الجمهورية العربية السورية

وزارة التعليم العالـي

جامعة البعث

كلية العلوم- قسم الرياضيات

استخدام الشبكات العصبونية في محاكاة المشاكل

USING NEURAL NETWORKS IN PROBLEMS SIMULATION

رسالة مقدمة لكلية العلوم في جامعة البعث للحصول على درجة

الماجستير في المعلوماتية

تقديم

وائل عبد السلام الحاج يونس

إشراف

الدكتورة رانية لطفي — كلية الهندسة المعلوماتية - جامعة البعث

الدكتور ناصر أبو صالح — كلية الهندسة المعلوماتية - جامعة البعث

2010م - 1431 هـ

الجمهورية العربية السورية

وزارة التعليم العالي

جامعة البعث

كلية العلوم – قسم الرياضيات

الأستاذ الدكتور عميد كلية العلوم المحترم

بعد اطلاع لجنة الحكم المشكلة لمناقشة رسالة الماجستير المقدمة من الطالب وائل الحاج يونس و التي هي بعنوان:

"استخدام الشبكات العصبونية في محاكاة المشاكل"

و المؤلفة من السادة الدكاترة:

الاسم و النسبة	المرتبة	الاختصاص	الكلية	الجامعة	الصفة
د.هلال هلال	أستاذ مساعد	الالكترونيات الدقيقة	هـــم.ك	حلبْ	رئيساً
د.يسر الأتاسي	أستاذ مساعد	ذكاء صنعي و بيانات	الهندسة المعلوماتية	البعث	عضواً
د.رانيا لطفي	مدرّس	ذكاء صنعي	الهندسة المعلوماتية	البعث	عضواً ومشرفاً

نحيطكم علماً بأن الطالب وائل عبد السلام الحاج يونس قد قام بإنجاز التوصيات العلمية و الفنية التي أوصت بها لجنة الحكم.

نرجو الإطلاع و اتخاذ ما ترونه مناسباً.

و لكم فائق الاحترام و التقدير

لجنة الحكم:

د. هلال هلال | د.يسر الأتاسي | د.رانيا لطفي

الفهارس

ملخص الأطروحة

في الآونة الأخيرة ازدادت الصعوبات التقنية و أصبحت أكثر تعقيداً و ما عاد للحلول البرمجية التقليدية القدرة على استيعابها، مما قاد الباحثين للتعمق في فهم آلية التفكير البشري و كيفية معالجته للمعلومات و تخزينها و استرجاعها و الاعتماد على أسلوب المحاكاة في حل هذه المشاكل فتم التوصل إلى هيكلية مترابطة لبنية برمجية سميت بالشبكات العصبونية الاصطناعية.

و هذا البحث يوضح قدرة الشبكات العصبونية في محاكاة المشاكل في مختلف المجالات و منها مجال معالجة الصور الرقمية و ذلك من خلال مقارنة النتائج مع طرق المعالجة التقليدية.

يتضمن البحث بناء ثلاث أنظمة معتمدة على الشبكات العصبونية للتعرف على الوجوه البشرية في الصور الملونة . تختلف الطرق الثلاث في نوع الدخل المقدم للشبكة. و تركز النقاش على شبكات معتمدة على :1)الصور المحرفة 2) مخطط RGB (RED – Green - Blue) الإحصائي 3) مخطط YES الإحصائي . و دُرب كل نظام على مجموعة الصور ذاتها باستخدام خوارزمية Levenberg-Marquad . وقُدمت النتائج الحسابية لمجموعة التدريب و مجموعة صور الاختبار الشائعة.

كما تضمن البحث تقديم طريقة تكيّفيّة لضغط الصور تعتمد على مستوى تعقيد الصورة. و كانت البنية الأساسية لأداة الضغط / فك الضغط في هذه الطريقة هي شبكة perceptron العصبونية متعددة الطبقات. و تُستخدم مختلف شبكات التوليد الخلفي (back-propagation) العصبية بالطريقة التكيّفيّة كأداة ضغط / و فك ضغط و ذلك بتقسيم الصورة إلى كتل و حساب التعقيد لكل كتلة و من ثم اختيار شبكة واحدة لكل كتلة تبعاً لدرجة تعقيدها .حيث يتم تقييم ثلاث طرق لقياس التعقيد و مقارنتها و هذه الطرق هي الفوضى أو الإنتروبية (Entropy) و الفعالية و المعتمدة على النمط و التي تُستعمل في تحديد مستوى التعقيد في كتل الصورة .و يُخصص لكل كتلة صورة شبكة عصبونية تبعاً لتعقيدها و ذلك في التدريب و التقييم .و أن

أفضل SNR (signal-to-noise Ratio) أو نسبة إشارة / تشويش هي البديل الآخر في اختيار شبكة ضغط كتل الصورة، حيث يؤدي اختيار واحدة من الشبكات المدربة إلى أفضل SNR في ضغط كتلة صورة الدخل .فحصلنا على أفضل النتائج حسب تقييماتنا و ذلك عندما يمكن للكتل أن تتداخل و يتم اختيار شبكة الضغط اعتماداً على أفضل SNR .

و تضمن البحث أيضاً طريقة لإزالة التشويش من الصور المشوهة. حيث تتضمن هذه الطريقة مرحلتين، المرحلة الأولى هي كشف التشويش، حيث نستخدم شبكة عصبونية من نوع هوبفيلد (Hopfield) لكشف تشويش البكسلات. و المرحلة الثانية هي مرحلة تصفية التشويش، حيث تم معالجة العيب الناتج من تطبيق متجه المتوسط(Median Vector) في فضاء وحيد اللون وتم الاعتماد على خوارزمية تقوم بتطبيق متجه المتوسط في الفضاء RGB أولاً، و من ثم في الفضاء Hue saturation intensity)HSI أو كثافة إشباع الشكل) ثانياً .و أن تطبيق الطريقة المقترحة على مجموعة من صور الاختبار القياسية أظهر أنها قادرة على اكتشاف التشويش الكلي في بكسلات الصورة مع نسبة خطأ ايجابي قريبة من الصفر، و إزالة التشويش، والاحتفاظ بأدق تفاصيل الصورة في نفس الوقت.

و قد تضمنت فصول البحث أيضاً مقدمة نظرية عن الشبكات العصبونية و بخاصة المستخدمة في هذا البحث و طرق تعلمّها و خوارزمية تدريبها و الآفاق المستقبلية.

مقدمة

تمهيد:

شهد العالم مؤخراً تطوراً ملحوظاً في كل ميادين الحياة ، و الكثير من هذه الميادين طرق بابها الحاسوب و على رأسها التحكم الصناعي و التطوير الطبي و كذلك قضايا إيجاد الحلول المثلى بالإضافة إلى أنواع أخرى من هذه الميادين و التي تتطلب مجالات التعرف على البصمات و الصور و الأشكال و التي تستخدم في الأبحاث الجنائية و غيرها الكثير.

إن حل المشكلة باستخدام الشبكات العصبونية، و اختيار الشبكة المناسبة للعديد من التطبيقات يتطلب الفهم المفصل لهذا النوع من الشبكات و التي تعتبر نتيجة لإنجاز العمليات الحسابية اليدوية و فهم طريقة تشفير بيانات المسألة على الحاسوب.استمر العمل بطريقة الحسابات اليدوية ، و لكن الطريقة التتابعية المستخدمة في الحسابات اليدوية و محدوديتها في حل العديد من المسائل الضخمة دفع العلماء للاستمرار في إيجاد طرق جديدة لمتابعة البحث في مجال الشبكات العصبونية الاصطناعية.

الهدف من البحث:

الهدف الأول:

الصور معرضة لعدة أنواع من التشويش (noise) . يوجد عدة طرق لإدخال التشويش إلى الصورة فعلى سبيل المثال :

أ- إذا كانت الصورة مكتسبة بشكل مباشر من صيغة رقمية ، فإن تقنية تأليف البيانات (مثل مكتشف أداة الشحن المزدوج charge coupled device أو CCD) يمكن أن يدخل تشويش .

ب-التحويل الالكتروني لبيانات الصورة يمكن أن يدخل تشويش .

فالهدف هنا تصميم مصفي (فلتر) للصور باستخدام الشبكات العصبية والتي تقوم بمسح كامل للصور المشوشة وتصفيتها من التشويش .

الهدف الثاني:

إن التعرف على الوجوه هي تمييز وجود و تحديد مكان وجه بشري في صورة ما.و لها تطبيقات هامة في الأنظمة الأمنية الآلية.

فالهدف هنا استخدام الشبكات العصبونية في التعرف على الوجوه من الصور الملونة.

الهدف الثالث:

تعتبر الصور الرقمية بيانات مركزة و بالتالي تحتاج إلى مساحات تخزين كبيرة. و نتيجة لذلك فإن إرسال صورة من جهاز لآخر قد يستهلك وقتاً كبيراً . و قد يؤدي استخدام تقنيات ضغط البيانات إلى إزالة بعض المعلومات من الصورة و لكن الضغط في هذه الحالة يوفر في مساحة التخزين و وقت الإرسال. للحفاظ على جودة الصورة و توفير مساحة التخزين و الوقت و على هذا يمكن استخدام الشبكات العصبونية في ضغط الصور.

الفصل الأول

مدخل إلى الشبكات العصبونية الاصطناعية

1-1 ما هي الشبكات العصبونية What is Neural Networks

1-1-1 الشبكات العصبونية الاصطناعية Artificial Neural Networks :

الشبكات العصبونية الصناعية هي عبارة عن نظام لمعالجة البيانات بشكل يحاكي و يشابه الطريقة التي تقوم بها الشبكات العصبية الطبيعية للإنسان أو للكائن الحي (أي النظام العصبي البشري) [1].

لقد تطورت الشبكات العصبونية كأمثلة رياضية معتمدة على طريقة التفكير البشري و كيفية معالجة الأعصاب للمعلومات. و الخطوات التالية تبين مكونات الشبكة العصبونية الاصطناعية و كيفية معالجتها للمعلومات:

1- تتم معالجة المعلومات في عناصر معالجة بسيطة تدعى العصبونات.

2- تمر الإشارات بين العصبونات عبر خطوط ربط.

3- يرفق كل خط ربط بوزن معين(قيمة عددية) و الذي يضرب مع الإشارات الداخلة إلى العصبون.

4- يطبق على كل عصبون تابع تنشيط ليحدد إشارة الخرج الناجمة عنه [2].

توصف الشبكات العصبونية بما يلي :

1- شكل الترابط بين العصبونات(المعمارية)

2- الطريقة التي تحدد الأوزان لهذه الترابطات(التدريب – التعليم – الخوارزمية).

3- نوع تابع التنشيط المستخدم.[3]

1-1-2 مكونات الشبكات العصبونية Neural Networks Components:

تتألف الشبكات العصبونية من عدد كبير من عناصر معالجة بسيطة تدعى العصبونات، كل عصبون يرتبط مع العصبونات الأخرى بأداة ربط مباشرة، و كل أداة تُرفق بوزن مُعيّن ، و هذه الأوزان تُمثل المعلومات التي ستبدأ بها الشبكة حل المشكلة [2].

إن لكل عصبون حالة داخلية خاصة به تدعى الفعالية، و التي حصل عليها من تطبيق تابع رياضي مُعين على معطيات الدخل و تم استقبال نتيجته من قبل العصبون.

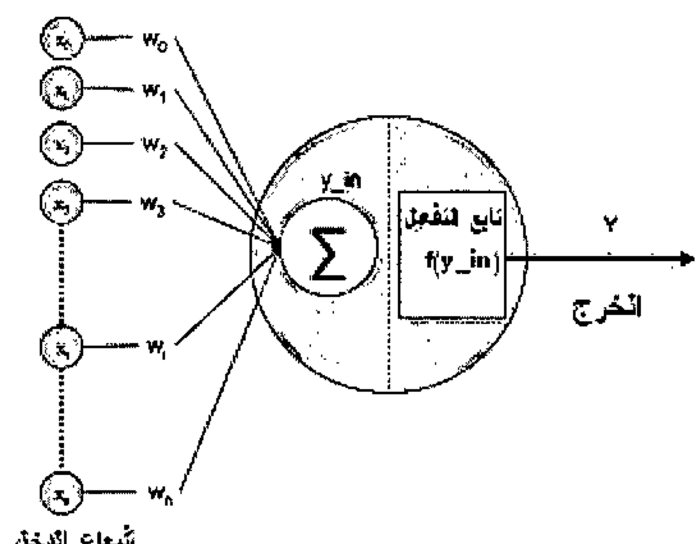
يُرسل كل عصبون فعاليته كإشارة دخل للعديد من العصبونات الأخرى المرتبطة معه.و بفرض لدينا العصبون Y الذي يستقبل دخله من العصبونات $X_1,X_2,X_3,...$ ، و التي فعاليتها(إشارات خرجها) هي x_1,x_2,x_3 على الترتيب (الشكل(1,1))، و لنرمز أيضاً بـــ $W_1,W_2,W_3,...$ لأوزان الترابطات القادمة من العصبونات $X_1,X_2,X_3,...$ إلى العصبون Y على الترتيب.إن دخل العصبون Y هو (y_in) و ناتج مجموع جداء كل إشارة دخل مع الوزن المرفق بها و العلاقة التالية توضح ذلك:

$$y_in = w_1.x_1 + w_2.x_2 + w_3.x_3 + \dots \qquad (1,1)$$

أما فعالية (إشارة خرج) العصبون Y تعطى بالعلاقة التالية:

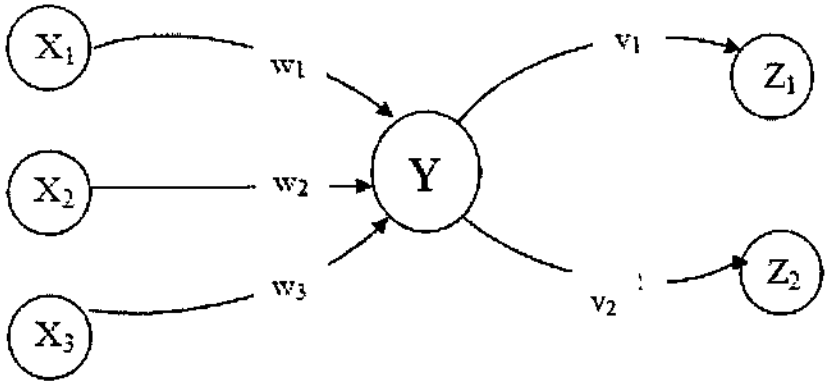
$$y = f(y_in) \qquad (2,1)$$

حيث f تابع رياضي معين يدعى تابع التفعيل[2].



الشكل(1-1) بنية العصبون

و لنفرض أن العصبون Y ارتبط مع عصبونات أخرى مثل Z_1,Z_2 حيث أوزان الارتباط الخاصة بها هي V_1,V_2 على الترتيب كما في الشكل(1-2). حيث يُرسل العصبون Y الإشارة y لكلا العصبونين Z_1,Z_2.



الشكل(1-2) البنية البسيطة للشبكة العصبونية

إن الوحدة Y تدعى بالطبقة المخفية و هذه الطبقة بالإضافة إلى تابع التفعيل تُعطي إمكانية أكبر في حل العديد من المشاكل و التي لا يمكن أن تُحل باستخدام وحدات دخل أو خرج فقط[1].

3-1-1 الشبكات العصبونية الحيوية Biological Neural Networks:

تتكون الخلية العصبية الطبيعية كما في الشكل (1-3) من الأجزاء التالية:

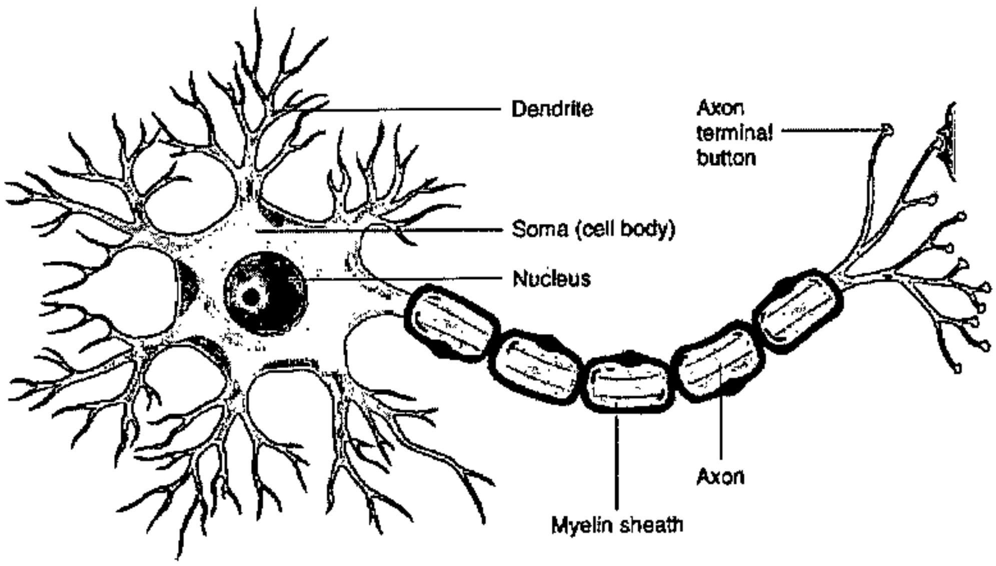
الجزء الأول Dendrites:

وهي عبارة عن متحسسات تقوم بالتقاط الإشارات العصبية من خلايا عصبية أخرى.ويمكن هنا أن نتخيل أن الخلية العصبية الطبيعية التقطت حرارة مرتفعة أو منخفضة فهنا تقوم مجموعة من خلايا الجلد للإنسان بتحويل العملية الكيميائية إلى إشارات عصبية يتم التقاطها من خلال المتحسسات.

الجزء الثاني Soma:

وهي تمثل جسم الخلية و هي تقوم على تجميع الإشارات المستقبلة من خلال المُتحسسات التي تُستخدم في المقارنة في الجزء Axon من الخلية.

الجزء الثالث Axon: وهي الجزء الذي يقرر أن يتم إرسال إشارة إلى الخلايا التي تلي الخلية الحالية [4].



الشكل (1−3) الخلية العصبية الطبيعية

إن المميزات الأساسية لعناصر المعالجة في الشبكات العصبونية الاصطناعية تُقترح اعتماداً على خصائص العصبونات الحيوية. و الخطوات التالية تبين تسلسل انتقال الإشارة ومعالجتها بين عناصر المعالجة:

1- تستقبل عناصر المعالجة العديد من الإشارات من الوسط الخارجي.

2- يُمكن أن تُعدل الإشارات الداخلة للعصبون عن طريق الأوزان في عقدة الاستقبال.

3- تجمع عناصر المعالجة أوزان الدخل القادمة لها من عصبونات أخرى.

4- تقوم العصبونات عند شروط معينة بإعطاء إشارة على خرجها.

5- إن الخرج الناتج عن هذه العصبونات يُمكن أن يُبث إلى عدة عصبونات أخرى [2].

2-1 كيفية استخدام الشبكات العصبونية

How Are Neural Networks Used

1−2−1 البنية الهندسية للشبكة Neural Architecture:

البنية الهندسية أو المعمارية للشبكة العصبونية الاصطناعية، هي الطريقة التي ترتبط بها العصبونات مع بعضها البعض لتشكيل الشبكة [4].يُمكن أن نتصور العصبونات المكوّنة

للشبكة مرتبة في طبقات، و تسلك جميع العصبونات الموجودة في نفس الطبقة طريقة واحدة في معالجتها للإشارات، إن العوامل الأساسية في تحديد سلوك وأداء العصبونات هي تابع التفعيل بالإضافة لعينة الترابطات الموزونة التي بالاعتماد عليها يتمكن العصبون من إرسال و استقبال الإشارات. وغالباً تمتلك جميع العصبونات في الشبكة نفس تابع التفعيل و نفس نوع الترابطات مع العصبونات الأخرى [2].

يُمكن أن ترتبط جميع العصبونات في نفس الطبقة مع عصبونات الطبقة الأخرى أو يُمكن أن تكون مرتبطة معها على الإطلاق[1]. و يُعتبر التفعيل لطبقة الدخل في الشبكات العصبونية هو إشارة الدخل القادمة من الوسط الخارجي، و يُوضح الشكل (1-2) أعلاه معمارية شبكة عصبونية نموذجية تتألف من طبقة دخل و طبقة خرج، و طبقة متوسطة بينهما هي الطبقة المخفية.[2]

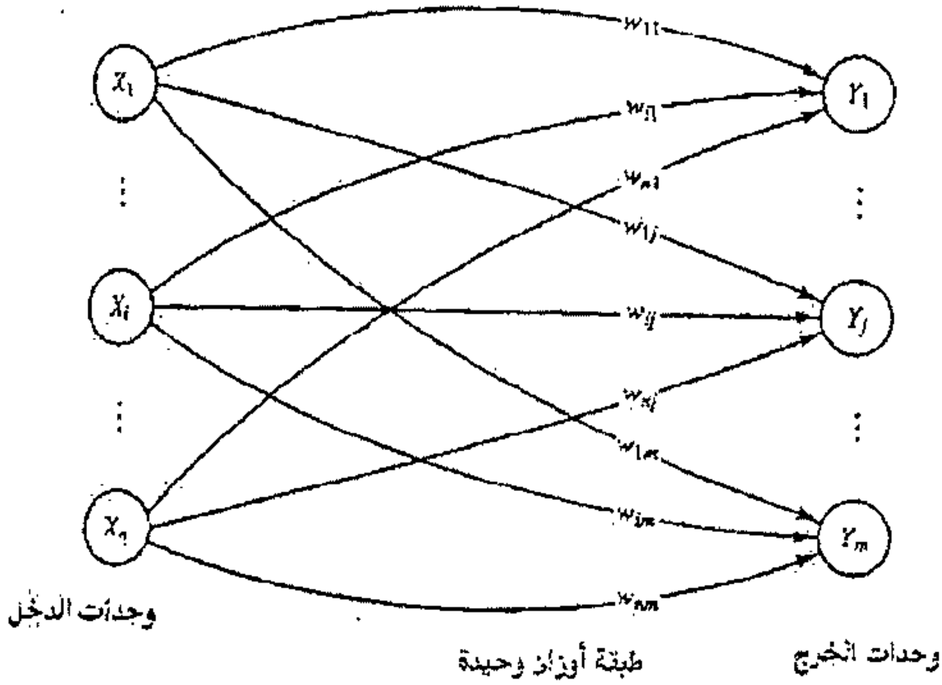
و تُصنف الشبكات العصبونية حسب عدد طبقاتها إلى:

1- شبكات وحيدة الطبقة (لا تملك طبقة مخفية).
2- شبكات متعددة الطبقة (تملك طبقة مخفية أو أكثر).

تتضمن أوزان الترابطات المعلومات التي ستتعلم الشبكة على أساسها.و يوضح الشكل (1-4) نموذج لشبكة عصبونية وحيدة الطبقة ذات انتشار أمامي للإشارة. تتدفق الإشارة في هذه الشبكات من وحدات الدخل باتجاه وحدات الخرج [2].

1-2-1-1 الشبكات وحيدة الطبقة Single Layer Networks:

يوجد طبقة واحدة من ترابطات الأوزان في هذا النوع من الشبكات. و تتميز هذه الشبكة بوجود طبقة دخل تستقبل الإشارة و طبقة الخرج التي نحصل منها على استجابة الشبكة. و يجب الملاحظة أنه لا يوجد ترابط بين قيم الأوزان بالنسبة لكل ترابط بين وحدات الدخل و وحدات الخرج.و تستخدم معمارية هذا النوع من الشبكات لحل المسائل المتعلقة بترابط العينات. و تُعتبر جميع العينات الناتجة عن استجابة الخرج هي مشاركة للعينات المقدمة من قبل وحدات الدخل.



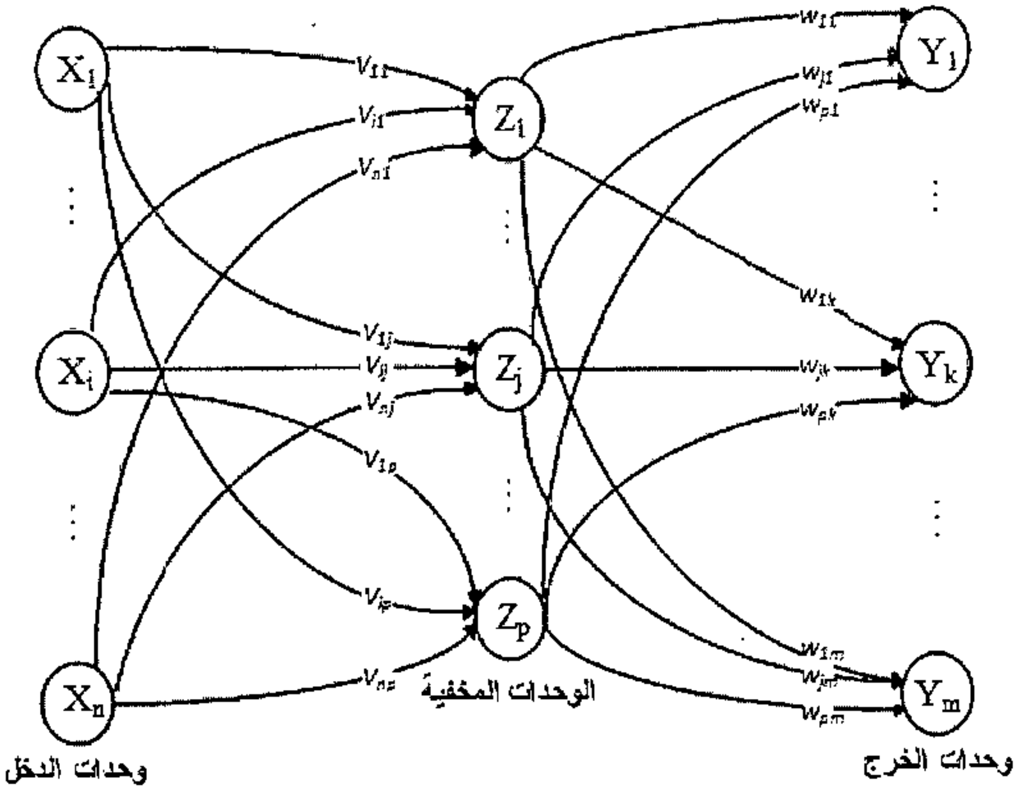
الشكل (1−4) شبكة عصبونية بطبقة وحيدة

و الفكرة الأساسية من تصنيف و ترابط العينات هو أنه باستخدام نفس معمارية الشبكة يُمكن حل مسائل مختلفة بالاعتماد على استقلالية الخرج الناتج عن الشبكة.

1−2−1−2 الشبكات متعددة الطبقة Multilayer Networks:

تتكون هذه الشبكات كما في الشكل (1−5) من طبقة واحدة أو أكثر من العصبونات و التي تدعى بالوحدات المخفية و تتوضع بين وحدات الدخل و وحدات الخرج. و يوجد بين كل طبقتين متجاورتين طبقة من ترابطات الأوزان(دخل −خرج)، (دخل − طبقة مخفية)، (طبقة مخفية − خرج) بالنسبة لشبكة تحوي طبقة مخفية وحدة مابين طبقتي الدخل و الخرج[2].

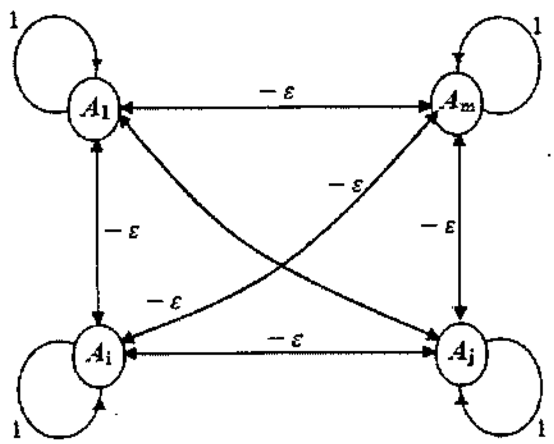
تستطيع الشبكات متعددة الطبقة مقارنةً مع الشبكات وحيدة الطبقة حل العديد من المشاكل المعقدة، ولكن تدريب هذه الشبكات يستغرق وقتاً أطول. حيث يمكن أن نتعرض لمشكلة لا يمكن حلها على الإطلاق باستخدام شبكة وحيدة الطبقة حتى لو دُربت لفترة طويلة [5].



الشكل (1-5) شبكة عصبية متعددة الطبقة

1-2-1-3 الطبقة المنافسة(التزاحمية) Competitive Layer:

تُعتبر الشبكات العصبونية ذات الطبقات المنافسة أحد أنواع الشبكات العصبونية المستخدمة في حل نوع خاص من المشاكل، و بشكل عام لتظهر التوصيلات الداخلية بين العصبونات في الطبقة التزاحمية ضمن مخططات معمارية هذه الشبكات.



الشكل (1-6) الطبقة التزاحمية

و يوضح الشكل (1-6) معمارية الطبقة التزاحمية و تمتلك التوصيلات الداخلية تملك قيم أوزان(ε -).

تتزاحم جميع التوصيلات الداخلية في الطبقة التزاحمية على الإشارة التي تُطبق عليها ليحصل عليها رابح واحد فقط [2].

1-2-2 توضع الأوزان Setting The Weights:

تُعتبر عملية توضع الأوزان خلال مرحلة التدريب من أهم المسائل و التي تُميز الشبكات العصبونية عن بعضها، إضافةً إلى معمارية الشبكة [2].

و تُقسم طرق تدريب الشبكات العصبونية إلى:

1- التدريب بمُعلم.

2- التدريب بدون مُعلم.

1-2-2-1 التدريب بمُعلم Supervised Training:

يتم في هذا النوع تدريب الشبكة على سلسلة من عينات تدريب الدخل التي يرافقها نموذج من عينات الخرج المرتبطة بها.

تُعدل الأوزان وفق خوارزمية تعليم تُدعى بالتدريب بمُعلم. و قد صُممت أول الشبكات معتمدة على هذا النوع من طرق التدريب في عملية تصنيف العينات. و يكون الخرج في هذه الشبكة عنصر معالجة قلاب أي: يأخذ الخرج القيمة (1) إذا كان شعاع الدخل تابعاً إلى أحد أصناف العينات، و يأخذ القيمة (1-) إذا لم يكن الشعاع تابعاً لأي صنف[2].

1-2-2-2 التدريب بدون مُعلم Unsupervised Training:

تُنظِّم الشبكات التي تعتمد مبدأ التدريب بدون مشرف أشعة الدخل المتشابهة بجانب بعضها تلقائياً بدون استخدام معطيات التدريب لتحدد ما هو العضو النموذجي من هذه المجموعة أو إلى أي مجموعة ينتمي هذا الشعاع. و يتم تقديم عدد من أشعة الدخل بدون تخصيص أي أشعة خرج مقابلة لها، و تستطيع هذه الشبكات تعديل أوزان ترابطاتها، و بالتالي

جميع أشعة الدخل المتشابهة تُوجه إلى نفس وحدة الخرج. و تُنتج هذه الشبكات أشعة مماثلة لأشعة التدريب، و هي نوع من شبكات التنظيم الذاتي للعالم Kohonen.

3-2-2-1 شبكات الأوزان الثابتة Fixed Weight Networks:

تستطيع العديد من الشبكات العصبونية إيجاد الحل الأمثل للمسائل المقيدة، وتستطيع شبكة الأوزان الثابتة حل هذا النوع من المسائل، و التي يصعب حلها بالطرق التقليدية. والتي من أجلها يجب أن تكون قيود المسألة مقبولة، وتستطيع الشبكة وفق هذه القيود الوصول إلى أقرب حل مثالي مقنع. و تُصمم هذه الشبكات بحيث تكون فيها أوزان الترابطات ثابتة لتمثيل القيود و كمياتها بحيث تكون الحلول المثلى أعظمية أو أصغرية.[2]

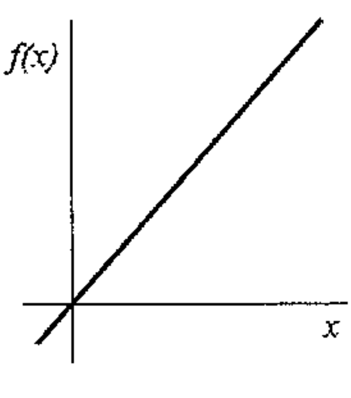
3-2-1 توابع التنشيط الشائعة Common Activation Functions:

تتضمن العملية الأساسية في خوارزميات التعلّم جمع إشارات الدخل المضروبة بالأوزان ثم تطبيق تابع التنشيط، ولا بد من استخدام توابع تنشيط معينة في خوارزميات التعليم[2].
و فيما يلي مجموعة من توابع التنشيط الشائعة المستخدمة في خوارزميات التعلم:

1- تابع التطابق:

$$f(x) = x \quad \text{For all } x \qquad (3\text{-}1)$$

و هو تابع خطي يُستخدم في الشبكات وحيدة الطبقة، و يُستخدم لتحويل دخل الشبكة إلى شكل مناسب لإشارة الخرج و هو قيمة متغيرة و مستمرة. و الشكل (1-7) يوضح هذا التابع.[2]

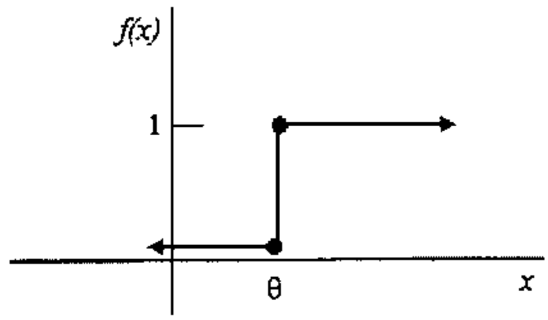


الشكل(1-7) تابع التطابق

2- تابع الخطوة الثنائي (ذو عتبة θ): وله الشكل:

$$f(x) = \begin{cases} 1 \ if \ x \geq \theta \\ 0 \ if \ x < \theta \end{cases} \qquad (4\text{-}1)$$

و يُوضح الشكل (1-8) هذا النوع من التوابع، و الذي يُعطي قيمة خرج ثنائية إما 0 أو 1، لذا يُدعى هذا التابع بتابع التنشيط الثنائي أو تابع الخطوة الثنائي، و يستخدم هذا التابع قيمة عتبة معينة تدعى θ[2] .



الشكل(1-8) تابع التنشيط الثنائي (ذو عتبة θ)

3- تابع سيغمويد الثنائي(Sigmoid):

و يوجد منه نوعين، هي التوابع النسبية و توابع مماس القطع الزائد، و لها مزايا خاصة أدت إلى استخدامها في خوارزميات تدريب الشبكات العصبونية ذات الانتشار العكسي. و توجد علاقة ربط بين هذه التوابع و مشتقاتها في نقطة معينة تخفف العبء عن الشبكة أثناء مرحلة التدريب من أجل معالجة الحسابات.و تقع قيمة هذه التوابع و مشتقاتها ضمن المجال[0,1] و غالباً ما تُستخدم كتوابع تفعيل من أجل الشبكات العصبونية التي تكون فيها قيمة الخرج هي قيمة ثنائية أو متغيرة ضمن المجال [0,1][2].

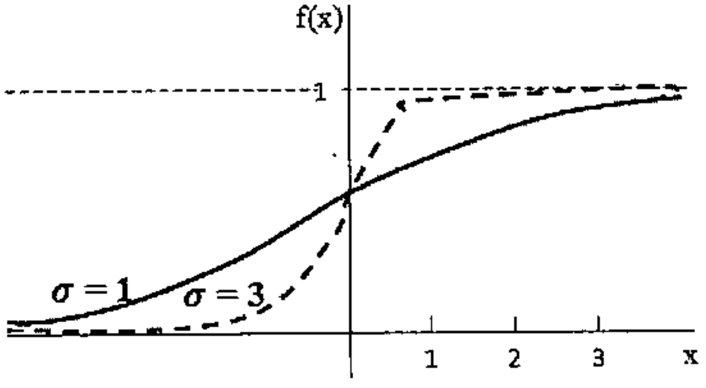
و الشكل الرياضي لهذه التوابع هو:

$$f(x) = \frac{1}{1+ e^{-\sigma x}} \qquad (5\text{-}1)$$

و مشتقه هو:

$$f' = \sigma . f(x). [1 - f(x)] \qquad (6\text{-}1)$$

يُوضح الشكل(1- 9) المخطط البياني لهذا التابع من أجل القيمتين $\sigma = 1$ و $\sigma = 3$.



الشكل (1−9) تابع تنشيط سيغمويد الثنائي($\sigma = 3, \sigma = 1$)

4-2-1 بعض الرموز المستخدمة Some of Used Symbols:

1- x_i, y_j: التنشيط الخاص بالوحدات X_i , Y_j على التوالي:

من أجل الدخل X_i فإن إشارة الدخل = x_i.و من أجل الوحدات الأخرى Y_j يكون :

$$y_j = f(y_in_j) \qquad (7\text{-}1)$$

2- W_{ij}: الوزن على الخط الواصل ما بين الوحدة X_i و الوحدة Y_j .

3- b_j الانحياز بالنسبة للوحدة Y_j حيث يعمل الانحياز عمل الوزن الرابط بين الوحدات و لكن تنشيطه ثابت يساوي الواحد.و الشكل (1−10) يوضح ذلك.

4- y_in_j: دخل الشبكة بالنسبة للوحدة Y_j و يتم حساب قيمته من العلاقة:

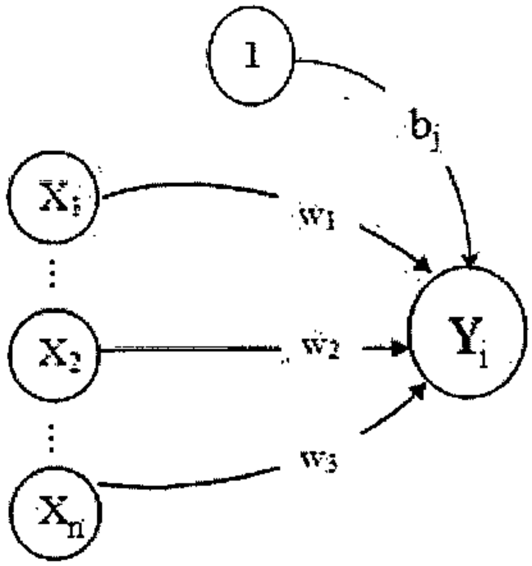
$$y_in_j = b_j + \sum x_i . w_{ij} \qquad (8\text{-}1)$$

5- W مصفوفة الوزن: و تأخذ الشكل التالي: $W = \{w_{ij}\}$.

6- W_{ij} شعاع الوزن :و يأخذ الشكل التالي: $W_{ij} = (w_{1j}, w_{2j}, \ldots, w_{nj})^T$ و هو العمود ذو الترتيب j من مصفوفة الأوزان.

7- $||X||$: القيمة المطلقة للشعاع X.

8- θ_j قيمة العتبة لتنشيط العصبون Y_j.



الشكل(1−10) بنية عصبون مع الانحياز

إن تابع تنشيط الخطوة يجعل قيمة تنشيط العصبون تأخذ القيمة(1) عندما تصبح قيمة دخل الشبكة أكبر من قيمة العتبة (θ_j). وخلاف ذلك فإن قيمة التنشيط تساوي(0) كما في الشكل(1−8).

9− S: شعاع دخل التدريب و يُعطى بالشكل: $S = (s_1,..,s_i,..s_n)$

10− t:شعاع خرج التدريب(الهدف) و يُعطى بالشكل: $t = (t_1,..,t_j,..s_m)$

11−X :شعاع الدخل و الذي من أجله ستستجيب الشبكة و يُعطى بالشكل:

$$X=(x_1,x_2,..,x_i,..,x_n)$$

12−Δw_{ij}: و هو التغير في الوزن W_{ij} و يعطى بالعلاقة:

$$\Delta w_{ij} = [w_{ij}\ (\text{new}) - w_{ij}\ (\text{old})] \qquad (9\text{-}1)$$

13− α معدل التعلم: يستخدم للتحكم في توليف الوزن خلال كل خطوة في مرحلة التدريب[2].

1−3 الشبكات العصبونية المستخدمة في تصنيف العينات:

Neural Networks for Pattern Classification

إن تصنيف العينات من المهام البسيطة التي يُمكن أن تُدرّب عليها الشبكات العصبونية لإظهار قدرتها و معالجتها للمعلومات [1].

يكون شعاع الدخل المقدم للشبكة في هذا النوع من المشاكل ينتمي أو لا ينتمي لصنف خاص أو ميزة خاصة[2].

1-3-1 شبكة Perceptron العصبونية:

تُستخدم شبكة Perceptron في التطبيقات التي تحتاج إلى تصنيف الأنماط إلى فئتين فقط. أي التي يمكن فصل الأنماط بها بيانياً بواسطة خط مستقيم وتسمى هذه النوعية من الأنماط: الأنماط القابلة للفصل خطياً[6].

وتتكون هذه الشبكة من طبقة الدخل(الوحدات الحسية) وطبقة واحدة أو طبقتين من طبقة المعالجة(وحدات الرابطة و وحدة الاستجابة) مشكلة نموذجاً يُشبه شبكية العين ، بالإضافة إلى أنه يتم ضبط الأوزان لطبقة واحدة فقط من طبقات الوصلات البينية التي تربط بين الطبقات السابقة، لتبقى الطبقة الأخرى (إن وجدت) ثابتة الأوزان. و تتعلم هذه الشبكة عن طريق معلم [6].

يتم اختيار تابع التنشيط الثنائي بالنسبة لوحدة الربط مع قيمة عتبة محددة.تُرسل الإشارة من وحدات الربط إلى وحدة الخرج على شكل إشارة ثنائية ممثلة بإحدى القيمتين (0 أو 1) و يُمثل الخرج بالعلاقة (y= f(y_in حيث يأخذ تابع التنشيط الشكل التالي:

$$f(y_in) = \begin{cases} 1 \;\; if \;\; y_in \; > \; \theta \\ 0 \;\; if \; -\theta \leq \; y_in \; \leq \; \theta \\ -1 \;\; if \;\; y_in \; < \; \theta \end{cases} \qquad (10-1)$$

تُعدل الأوزان القادمة من وحدات الربط إلى وحدة الخرج وفق قاعدة Perceptron في التعلّم.و من أجل كل عينة تدريب ستحسب الشبكة استجابتها، و عندها يتم تحديد الخطأ الذي ارتكبته الشبكة أثناء مرحلة التدريب بمقارنة قيمة الخرج المحسوب مع الخرج الفعلي للشبكة[2].

يتم تعديل الأوزان اعتماداً على الخطأ المرتكب حيث تُعدل الأوزان بين الترابطات القادمة من وحدات الربط غير الصفرية إلى وحدة الخرج إذا لم يكن الخرج المحسوب (0) و الخرج الفعلي(1−) أو إذا لم يكن الخرج المحسوب هو (1+) و الخرج الفعلي هو (1−).
و تُعطى العلاقة التي تُعبر عن عملية تعديل الأوزان بالشكل:

$$w_i(new) = w_i(old) + \alpha . t . x_i \qquad (11-1)$$

حيث قيمة الخرج الفعلي للشبكة (t) هي (+1) أو (-1)، α مُعدل التعلّم. و في حال عدم حدوث خطأ فالأوزان سوف لن تُعدل. تستمر الشبكة في تعلمها حتى تصبح قيم الخرج المحسوبة موافقة لقيم الخرج الفعلي.[1]

1-1-3-1 الخوارزمية Algorithm:

الخطوة1: تهيئة قيم الأوزان و الانحياز (قيم صفرية للتبسيط)، ضع معدل التعلم بقيمة $(0 < \alpha \leq 1)$(للتبسيط نجعل $\alpha = 1$).

الخطوة 2: طالما لم يتحقق شرط التوقف(عندما تصبح قيم الخرج المحسوبة موافقة لقيم الخرج الفعلي) نفذ الخطوات من 3 ⟵ 7 .

الخطوة 3: من أجل كل زوج تدريب (S:t) نفذ الخطوات من 4 ⟵ 6.

الخطوة 4:ضع قيم تنشيط وحدات الدخل بالشكل $x_i = s_i$.

الخطوة 5: احسب استجابة وحدة الخرج بالعلاقة:

$$y_in = b + \sum_i x_i . w_i$$

$$y = \begin{cases} 1 & if \quad y_in > \theta \\ 0 & if \quad -\theta \leq y_in \leq \theta \\ -1 & if \quad y_in < \theta \end{cases}$$

الخطوة 6: حسّن الأوزان و الانحياز في حالة حدوث خطأ ما لهذه العينة كما يلي:

إذا كان $y \neq t$ (الخرج المحسوب ≠ الخرج المرغوب) عندئذ:

$$w_i(new) = w_i(old) + \alpha . t . x_i$$
$$b(new) = b(old) + \alpha . t$$

و إلا:

$$w_i(new) = w_i(old)$$
$$b(new) = b(old)$$

الخطوة 7: اختبر شرط التوقف. إذا لم تتغير الأوزان التي تم تهيئتها في الخطوة 3 توقف و إلا أعد الخطوات من 3 ⟵ 7[1].

1 – 4 شبكة الترابط الذاتي Auto Associative Network :

تُعتبر شبكة الترابط الذاتي شبكة ذات انتشار أمامي. في هذه الشبكات تكون أشعة التدريب للدخل و الخرج متطابقة [2].

في هذا النوع من الشبكات، يتم تخزين أزواج أشعة الدخل و الخرج، و يتم اعتماد التمثيل الثنائي أو ثنائي القطبية، و يُمكن لهذه الشبكة استرجاع العينات المخزنة لديها كلياً أو جزئياً في حال تعرضت العينة المدخلة لبعض التشويش إذا كان الدخل المقدم مشابه إلى حد معين مع الدخل المخزن في الشبكة. و يتم الحكم على قدرة هذه الشبكة من حيث قابليتها في إعادة إنتاج العينات المخزنة لديها انطلاقاً من عينات اختبار مشوشة[2].

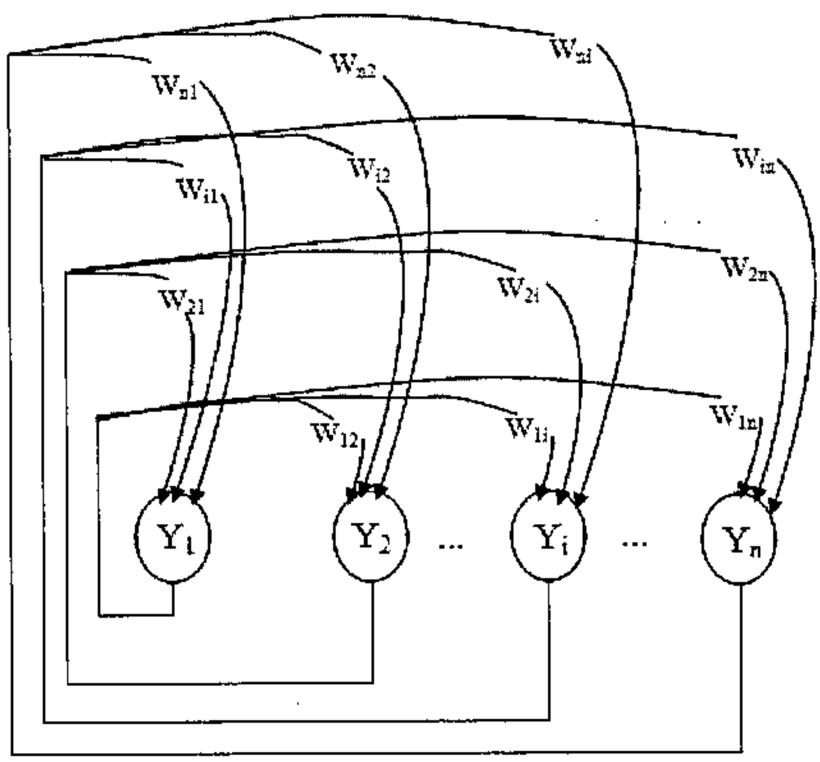
1-4-1 شبكة هوبفيلد Hopfield Network:

و هي عبارة عن شبكة عصبونية متصلة فيما بينها داخلياً، و كل وحدة فيها ترتبط مع جميع الوحدات الأخرى و تستخدم مصفوفة أوزان متناظرة. و بدون وصلات ذاتية كما في العلاقة: $W_{ii} = 0$ و $W_{ij} = W_{ji}$ و تتميز شبكة هوبفيلد بما يلي:

1- تقوم وحدة واحدة بتحسين تنشيطاتها في لحظة معينة معتمدة على الإشارة المستقبلة من كل وحدة أخرى.

2- كل وحدة تستمر في استقبال الإشارة الخارجية، بالإضافة إلى الإشارة القادمة من وحدات أخرى في الشبكة [2].

يوضح الشكل(1-11) الشكل العام لشبكة هوبفيلد.



الشكل (1-11)البنية الهندسية لشبكة هوبفيلد

1-4-1-1 الخوارزمية Algorithm:

الخطوة1:تهيئة الأوزان لتخزين العينات كما يلي:

$$W = \begin{bmatrix} 0 & 1 & 1 & -1 \\ 1 & 0 & 1 & -1 \\ 1 & 1 & 0 & -1 \\ -1 & -1 & -1 & 0 \end{bmatrix}$$

الخطوة2: من أجل كل شعاع دخل X ، نفذ الخطوات 3 ⟵ 7

الخطوة3: ضع تنشيطات البداية للشبكة مساوية لشعاع الدخل الخارجي X بالشكل:

$$y_i = x_i \quad ; \quad i = 1, \dots, n$$

الخطوة4: نفذ الخطوات من 5 ⟵ 7 من أجل كل وحدة y_i.

الخطوة5: احسب دخل وحدة الخرج بالعلاقة:

$$y_in_1 = x_1 + \sum_j y_j . w_{j1}$$

الخطوة6:احسب تنشيط وحدات الخرج بالعلاقة:

$$y = \begin{cases} 1 \ \ if \ \ y_in_i > \theta_i \\ y_i \ \ if \ \ y_in_i = \theta_i \\ 0 \ \ if \ \ y_in_i < \theta_i \end{cases}$$

الخطوة7: وزع القيمة y_i إلى جميع الوحدات الأخرى.

1-5 الشبكات العصبونية ذات الانتشار الخلفي

Backpropagation Neural Networks

يتم في هذه الشبكات استخدام مجموعة من أشعة الدخل لتحديد مجموعة من أشعة الخرج، و تستخدم هذه الشبكات أسلوب التعلم بمشرف[1].

تهدف عملية تدريب هذه الشبكات إلى الوصول إلى حالة من التوازن بين قابلية الشبكة على الاستجابة الصحيحة لعينات الدخل التي تستخدم في عملية التدريب(يتم تخزينها في الشبكة) و قابلية الشبكة على إعطاء استجابة جيدة لدخل مشابه لكن غير مطابق لذلك الدخل المستخدم في التدريب[2].

يتضمن تدريب الشبكة بطريقة الانتشار الخلفي (العكسي) ثلاثة مراحل:

1- مرحلة التغذية الأمامية لعينات تدريب الدخل.

2- مرحلة الحساب والانتشار الخلفي للخطأ المتعلق بالخرج.

3- مرحلة توليف الأوزان.

وبعد انتهاء مرحلة اختبار الشبكة و التي تتضمن مرحلة واحدة فقط هي مرحلة الانتشار الأمامي.

و حتى إذا كان التعليم بطيئاً، فإن الشبكة المتدربة يمكنها أن تنتج إشارة خرجها بشكل سريع جداً. لقد تم تطوير طريقة الانتشار العكسي و ذلك لتحسين سرعة العملية التدريبية.

1-5-1 الخوارزمية Algorithm:

بما أن تدريب الشبكة العصبونية باستخدام الانتشار الخلفي يتضمن ثلاث مراحل هي الانتشار الأمامي لعينة تدريب الدخل، و الانتشار الخلفي للخطأ، و توليف أوزان الشبكة.

خلال مرحلة الانتشار الأمامي: تستقبل كل وحدة دخل(X_i) إشارة الدخل و من ثم تنتشر هذه الإشارة إلى كل وحدة من وحدات الطبقة المخفية(Z_1 ,...,Z_p) ثم تحسب كل وحدة من وحدات الطبقة المخفية بدورها قيمة التنشيط لهذه الإشارة و بعدئذ ترسل هذه الوحدات إشارتها (Z_j) إلى كل وحدة من وحدات طبقة الخرج. ثم تقوم كل وحدة من وحدات طبقة الخرج (Y_k) بحساب تنشيطها(y_k) لتشكل استجابة الشبكة من أجل عينة الدخل المعطاة[2].

خلال مرحلة الانتشار الخلفي للخطأ: تقوم كل وحدة خرج بمقارنة تنشيطاتها المحسوبة y_k مع قيمة الخرج الفعلي t_k لتحدد قيمة الخطأ الحاصل لهذه العينة لتلك الوحدة. و اعتماداً على قيمة الخطأ الحاصل فإن العامل(δ_k يتم حيث $k = 1,\ldots,m$)يتم حسابه، حيث يُستخدم العامل δ_k لتوزيع الخطأ على وحدات طبقة الخرج Y_k لتتم إعادته إلى كل الوحدات في الطبقة السابقة(وحدات الطبقة المخفية المرتبطة مع وحدات الخرج Y_k). و كذلك يُستخدم هذا العامل فيما بعد لتحديث الأوزان بين طبقة الخرج و الطبقة المخفية. و بطريقة مشابهة، فإن العامل δ_j حيث ($j = 1,\ldots,p$) يتم حسابه بالنسبة لكل وحدة من وحدات الطبقة المخفية Z_j، و ليس ضرورياً لنشر الخطأ خلفاً إلى وحدات طبقة الدخل و لكن العامل δ_j يُستخدم لتحديث الأوزان بين الطبقة المخفية و طبقة الدخل[2].

الرموز المستخدمة في الخوارزمية:

X : شعاع تدريب الدخل $X = (x_1 , \ldots,x_i,\ldots,x_n)$

T : شعاع الهدف للخرج (المرغوب)$T = (t_1 , \dots , t_k , \dots ,t_m)$

δ_k : عامل تصحيح الخطأ لتوليف الوزن W_{jk} و الذي يكون عائداً لقيمة الخطأ في الخرج Y_k، و يقدم معلومات عن الخطأ في الوحدة Y_k و الذي يتم انتشاره عكسياً إلى وحدات الطبقة المخفية التي تتغذى من الوحدة Y_k.

δ_j:عامل تصحيح الخطأ لتوليف الوزن V_{ij} العائد لانتشار معلومات الخطأ من طبقة الخرج إلى الوحدة المخفية Z_j.

α:مُعدل التعلم.

X_i:وحدة الدخل ذات الترتيب i ، و من أجل وحدة الدخل فإن إشارة الدخل و إشارة الخرج تكون نفسها أيx_i.

V_{0j}: الانحياز على الوحدة المخفية ذات الدليل j>

Z_j:وحدة الطبقة المخفية ذات الدليل j ، حيث أن دخل الوحدات المخفية يرمز له z_in_j و يعطى بالعلاقة:

$$z_in_j = V_{0j} + \sum_i x_i v_{ij}$$

و خرج وحدات الطبقة المخفية Z_j (قيمة التنشيط) يرمز لها بالرمز z_j و يتم حسابها وفق العلاقة:

$$z_j = f(Z_in_j)$$

W_{ok}:الانحياز على وحدة الخرج ذات الدليل k.

Y_k:وحدة طبقة الخرج ذات الدليل K، حيث أن دخل هذه الوحدة يرمز له بالرمز y_in_k و يتم حسابه بالعلاقة:

$$y_in_k = w_{ok} + \sum_j z_j . w_{ik}$$

و خرج هذه الوحدة(تنشيطها) يُرمز لها بالرمز Y_k و يتم حسابها وفق العلاقة:

$$y_k = f(y_in_k)$$

خوارزمية التدريب Training Algorithm:

الخطوة1: تهيئة الأوزان(وضعها بقيم عشوائية صغيرة).

الخطوة2:طالما أن شرط التوقف لم يتحقق بعد، نفذ الخطوات من 3 ⟵ 10.

الخطوة3: من أجل كل زوج تدريب نفذ الخطوات من4 ⟵ 9.

مرحلة الانتشار الأمامي (التغذية بالاتجاه الأمامي):

الخطوة4: تستقبل كل وحدة دخل(X_{j}, i = 1 , … , n) إشارة دخلها x_j ثم تنشرها إلى جميع وحدات الطبقة المخفية.

الخطوة5:تجمع كل وحدة مخفية (Z_j , j = 1,… , p)قيم و إشارات دخلها الموزونة كما يلي:

$$z_in_j = v_{0j} + \sum_{i=1}^{n} x_i v_{ij}$$

بعد ذلك يطبق تابع تنشيطها و تحسب إشارة خرجها:

$$z_j = f(z_in_j)$$

و ترسل قيمة التنشيط إلى جميع وحدات الطبقة الأعلى(طبقة الخرج).

الخطوة6:تجمع كل وحدة خرج(Y_k ; k = 1 , … ,m) إشارات دخلها الموزونة كما يلي:

$$y_in_k = w_{0k} + \sum_{j=1}^{p} z_j . w_{jk}$$

بعد ذلك يطبق تابع التنشيط لحساب إشارة خرجها أي:

$$y_k = f(y_in_k)$$

مرحلة الانتشار الخلفي للخطأ:

الخطوة7:تستقبل كل وحدة من وحدات طبقة الخرج(Y_k) العينة الهدف(المرغوبة) المقابلة لعينة تدريب الدخل. بعدئذ يتم حساب حد معلومات الخطأ لها:

$$\delta_k = (t_k - y_k).\ f'(y_in_k)$$

ثم يتم حساب حد تصحيح أوزانها المستخدم لتحديث الوزن w_{jk} لاحقاً:

$$\Delta w_{jk} = \alpha . \delta_k . z_j$$

الخطوة8:تجمع كل وحدة في الطبقة المخفية إشارات δ من الوحدات في الطبقة الأعلى.

مرحلة تحديث الأوزان و الانحيازات:

الخطوة9:تحدث كل وحدة خرج انحيازها و أوزانها كما يلي:

$$w_{jk}\ (new) = w_{jk}(old) + \Delta w_{jk}$$

بالنسبة لكل وحدة مخفية حدّث الانحيازات و الأوزان:

$$V_{ij}(new) = V_{ij}(old) + \Delta V_{ij}$$

حيث: $\alpha . \delta_j . x_i = \Delta V_{ij}$

الخطوة10:اختبر شرط التوقف.

الفصل الثاني

التعرف على الوجوه باستخدام الشبكات العصبونية

1-2 مقدمة Introduction:

إن مشكلة التعرف على الوجوه هي تمييز وجود الوجه البشري في صورة ما.و لهذه المشكلة تطبيقات هامة في الأنظمة الأمنية الآلية و قراءة الشفاه و فهرسة و استرجاع صور الفيديو و المؤتمرات عبر الفيديو. هناك تقنيتان أساسيتان للتعرف على الوجوه :طرق معتمدة على المحتوى و طرق معتمدة على اللون.

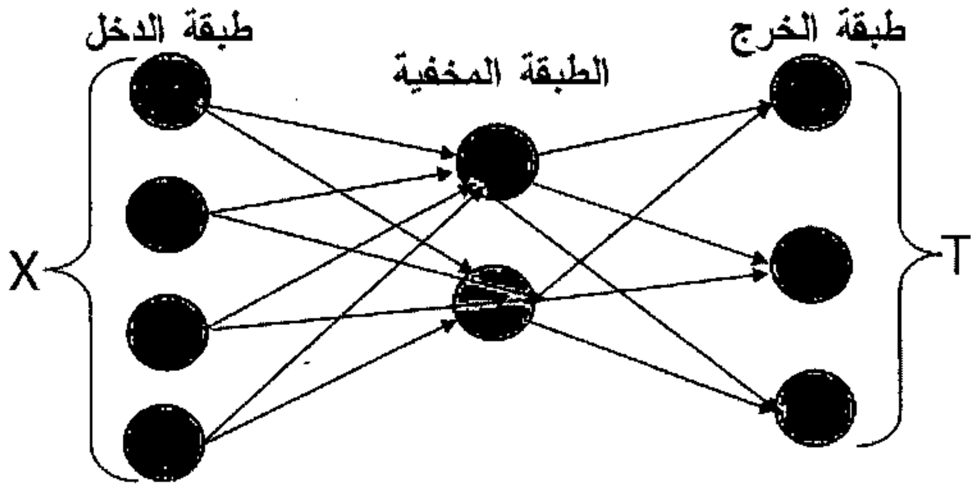
تحاول الطرق المعتمدة على المحتوى التعرّف على السمات في الوجه البشري . و قد طُورت معظم الطرق المعتمدة على المحتوى من أجل الصور ذات التدرج الرمادي لتفادي تعقيد دمج السمات المكتشفة في فضاء اللون RGB. قام ياو و كيبولا بتطوير طريقة تؤدي إلى استطالة الصورة أفقياً و تميز السمات الأفقية الرقيقة مثل العيون و الفم باستخدام مرشحات مشتق غوص من الدرجة الثانية[8]. و قام كوتس و تايلور بتطوير طريقة تطابق السمات مع وجه نموذجي باستعمال طرق إحصائية[9]. كما قدم بيرل و بيرونا طريقة مشابهة تطابق السمات مع وجه نموذجي و لكنهم استعملوا خوارزمية مطابقة بيانية لمقارنة السمات المكتشفة مع النموذج [10]. و طوّر رولي و بلوجا و كينيد نظام لكشف الوجه من الجهة الأمامية يستخدم الشبكات العصبية لاختيار السمات[11]. كما طوّر سونغ و بوجيو طريقة تعلم تعتمد على مثال دون استخدام الشبكات العصبونية[12]. بينما استعمل كومنزر و هوانغ نظام تعلّم بصري احتمالي [13-14]. ولكن التقنيات المعتمدة على المحتوى معقدة جداً و مكلفة بشكل حسابي.

تقوم التقنيات المعتمدة على اللون بحساب مخطط إحصائي لقيم اللون و من ثم تطوير مخطط كثافة اللون لتمييز احتمال تمثيل مجالٍ معين من قيم البكسلات لجلد بشري. و قد وجد أن فعالية الطريقة تعتمد إلى حد كبير على فضاء اللون المستعمل. فقد تم تطوير مخططات كثافة اللون من أجل فضاء اللون RGB القياسي[15]، و فضاء اللون YIQ [16]، و فضاء اللون HSV [17]، و الفضاء LUV [18].إن تطبيق التقنيات المعتمدة على اللون بسيطة جداً و يصبح النظام كفؤ جداً بعد تعلّم مخططات كثافة اللون.و لكن يوجد للتقنيات المعتمدة على اللون عدة عوائق[19].و تتضمن هذه العوائق فقدان المعلومات بسبب تجزئة الصورة و الاعتماد الكبير على فضاء اللون.و لكن العائق الأهم هو التقنية المعتمدة فقط على مدرج اللون الإحصائي فقط مع تجاهل المعلومات المكانية للصورة. وعلى الرغم من تلك السلبيات فإن مخططات اللون الإحصائية شائعة جداً بسبب بساطتها و سهولتها في الحساب.

نريد في هذه الطريقة حل مشكلة التعرف على الوجوه البشرية بأسلوب قريب من الطريقة التي يتعلم بها الدماغ البشري التصنيف. و نحن نريد من النظام (الشبكة العصبونية)أن يتعلم التمييز بين الوجه البشري و سواه.سنناقش الشبكات العصبية في المقطع الثاني بإيجاز و التعلّم بدون إشراف أو التدريب المستخدم لوضع أوزان الشبكة.و لتطوير النظام سنتابع بأسلوب تجريبي تغيير أنواع الدخل المقدم للشبكة العصبية. في المقطع الثالث سنقدم طريقة لتغذية الشبكة العصبية بقيم RGB للبكسلات المباعدة بانتظام في صورة ملونة. و كان المطلوب من الشبكة أن تتعلّم المعلومات المكانية ثم تقوم بإعطاء القرار و لكن هذه التقنية المعتمدة على المحتوى كما ناقشنا أعلاه مكلفة جداً بشكل حسابي و النتائج كانت ليس كما يجب أن تكون عليه.في المقطع الرابع نناقش الطريقة الثانية التي تعتمد على إنشاء مخطط اللون الإحصائي في فضاء اللون RGB. و تكمن الفكرة في تعلم الشبكة العصبية مخطط كثافة لون من مجموعة التدريب. تدريب الشبكة كان سهل بشكل حسابي و لكن النتائج كانت متفاوتة. في المقطع الخامس نصف الطريقة الثالثة حيث تم تدريب الشبكة العصبية على مخططات اللون الإحصائية في فضاء اللون YES.فكانت الكلف الحسابية مماثلة للكلف في فضاء اللون RGB و لكن نتائج الاختبار كانت أفضل بكثير. استخدمنا مجموعة صور التدريب و الاختبار ذاتها لنصبح قادرين على المقارنة بين الطرق الثلاث. وأخيراً في خلاصتنا في المقطع السادس نناقش محاسن كل طريقة و المجالات الممكنة لتحسينها.

2-2 الشبكات العصبونية و التعرف على الوجوه:

Neural Networks and Face recognition



الشكل(2−1) مخطط الشبكة العصبية العام و المستخدم في التعرف على الوجوه

تشمل مجموعة معطيات التدريب مجموعة أشعة الدخل X مع أشعة الخرج المطابقة T. و يتضمن التدريب الفعلي وضع الأوزان و بالتالي يحسب شعاع الخرج T_{NET} بواسطة الشبكة من أجل كل دخل X ليطابق الخرج المطلوب T_{ACTUAL} بأكبر قدر ممكن و نحن نريد إيجاد مجموع الأوزان الذي يجعل $\|T_{NET} - T_{ACTUAL}\|$ أقل ما يمكن. و باختيار مجموعة التدريب بعناية لتمثل كامل فضاء أشعة الدخل المحتملة فإن الدخل المشابه لأخر في مجموعة التدريب يجب أن يعطي خرج مشابه.إن التدريب مكلف جداً حسابياً و لكن يجب تحديد الأوزان لكل المعطيات في مجموعة التدريب. وعلى كل حال بعد انتهاء التدريب فإن حساب الخرج T من أجل الدخل X يظهر فعالية عالية.

بالنسبة لمشكلة التعرف على الوجه سوف يتضمن شعاع الدخل X المعلومات المشتقة من الصورة الملونة. و سيكون شعاع الخرج T عبارة عن عدد مفرد يمثل احتمال احتواء الصورة على وجه بشري. فإذا جعلنا نمط ωليكون وجهاً بشرياً و ملاحظة صورة ملونة x، و من ثم نحاول حساب $P(\omega|x)$ (أي احتمال احتواء الصورة x على النمط ω) أو للتبسيط P. و يجب أن نلاحظ بأن تفسير P كاحتمال قد لا يتضمن خرج الشبكة الفعلي، لكون أنه ليس هناك ضمانة بأن كل دخل سيؤدي إلى خرج P حيث $0 \le P \le 1$. لذا نفسر الخرج P لصورة الدخل المعطاة X كما يلي:

$$P=\begin{cases} > 0.5 \implies X \text{ يحتوي وجه بشري} \\ < 0.5 \implies X \text{ لا يحتوي وجه بشري} \\ = 0.5 \Rightarrow \text{غير واضح إذا كان} X \text{ يحتوي وجه بشري} \end{cases}$$

نختار شبكة عصبية من النمط الأمامي التغذية Feed-Forward لها طبقة مخفية واحدة (لأننا سنعالج نمط واحد من المعطيات ألا و هو الصورة) تحتوي على 20 عقدة (بعدد صور التدريب). إذا يكون عدد الأوزان $20(|X|+1)$. نختار دالة sigmoid $\sigma(t)$ كدالة عتبة حيث:

$$\sigma(t) = \frac{1}{1+e^{-t}}$$

خوارزمية التدريب التي كانت مستعملة هي خوارزمية Levenberg-Marquad[19]، و التي تقلل الخطأ $\|T_{NET} - T_{ACTUAL}\|$ (الفرق بين الخرج المطلوب و الخرج الفعلي)[13] و هي تحقق نسبة تقارب أفضل و تتطلب ذاكرة أقل و أسرع من غيرها من خوارزميات تدريب الشبكات العصبونية الأخرى مثل خوارزمية الانتشار المرتد بالخطأ (Error Back propagation Algorithm) حيث تكون هذه الخوارزمية تكون غالباً بطيئة في المشاكل العملية حيث يمكن أن تكون خوارزمية Levenberg-Marquad أسرع منها من 10 حتى

100 مرة وهي تعتبر تطوير لخوارزمية الانتشار المرتد بالخطأ حيث أن مؤشر الأداء الذي سيتم تحسينه بواسطة خوارزمية Levenberg-Marquad معرف كما يلي:

$$F(w) = \sum_{p=1}^{P}\left[\sum_{k=1}^{K}\left(d_{kp} - o_{kp}\right)^2\right] \qquad (1-2)$$

حيث $w = [w_1 \ \ w_2 \ \ \ldots \ \ w_N]^T$ يتضمن كل أوزان الشبكة، d_{kp} القيمة المطلوبة من الخرج k و النمط p و o_{kp} هو القيمة الفعلية للخرج k و النمط p و N هو عدد الأوزان وP عدد الأنماط و K هو العدد الكلي لخرج الشبكة. العلاقة (2−1) يمكن أن تكتب كما يلي :

$$F(w) = E^T E \qquad (2-2)$$

حيث:

$$E = [\ e_{11} \ \ldots \ e_{k1} \ \ e_{12} \ \ldots \ e_{k2} \ \ldots \ e_{1p} \ldots \ e_{kp}\ ]^T$$

$$e_{kp} = d_{kp} - o_{kp} \ , \ k = 1,\ldots,K \ , \ p = 1 \ \ldots \ P$$

حيث E هو شعاع الخطأ المتراكم (من أجل جميع الأنماط). يمكن أن نعرف المصفوفة اليعقوبية من العلاقة (2−2) كما يلي :

$$J = \begin{bmatrix} \frac{\partial e_{11}}{\partial w_1} & \frac{\partial e_{11}}{\partial w_2} & \cdots & \frac{\partial e_{11}}{\partial w_N} \\ \frac{\partial e_{21}}{\partial w_1} & \frac{\partial e_{21}}{\partial w_2} & \cdots & \frac{\partial e_{21}}{\partial w_N} \\ \vdots & \vdots & & \vdots \\ \frac{\partial e_{k1}}{\partial w_1} & \frac{\partial e_{k1}}{\partial w_2} & \cdots & \frac{\partial e_{k1}}{\partial w_N} \\ \vdots & \vdots & & \vdots \\ \frac{\partial e_{1p}}{\partial w_1} & \frac{\partial e_{1p}}{\partial w_2} & \cdots & \frac{\partial e_{1p}}{\partial w_N} \\ \frac{\partial e_{2p}}{\partial w_1} & \frac{\partial e_{2p}}{\partial w_2} & \cdots & \frac{\partial e_{2p}}{\partial w_N} \\ \vdots & \vdots & & \vdots \\ \frac{\partial e_{kp}}{\partial w_1} & \frac{\partial e_{kp}}{\partial w_2} & \cdots & \frac{\partial e_{kp}}{\partial w_N} \end{bmatrix} \qquad (3-2)$$

و تحسب الأوزان بالعلاقة التالية:

$$\mathbf{w_{t+1} = w_t - (J_t^T J_t + \boldsymbol{\mu} I)^{-1} J_t^T E_t} \qquad (4-2)$$

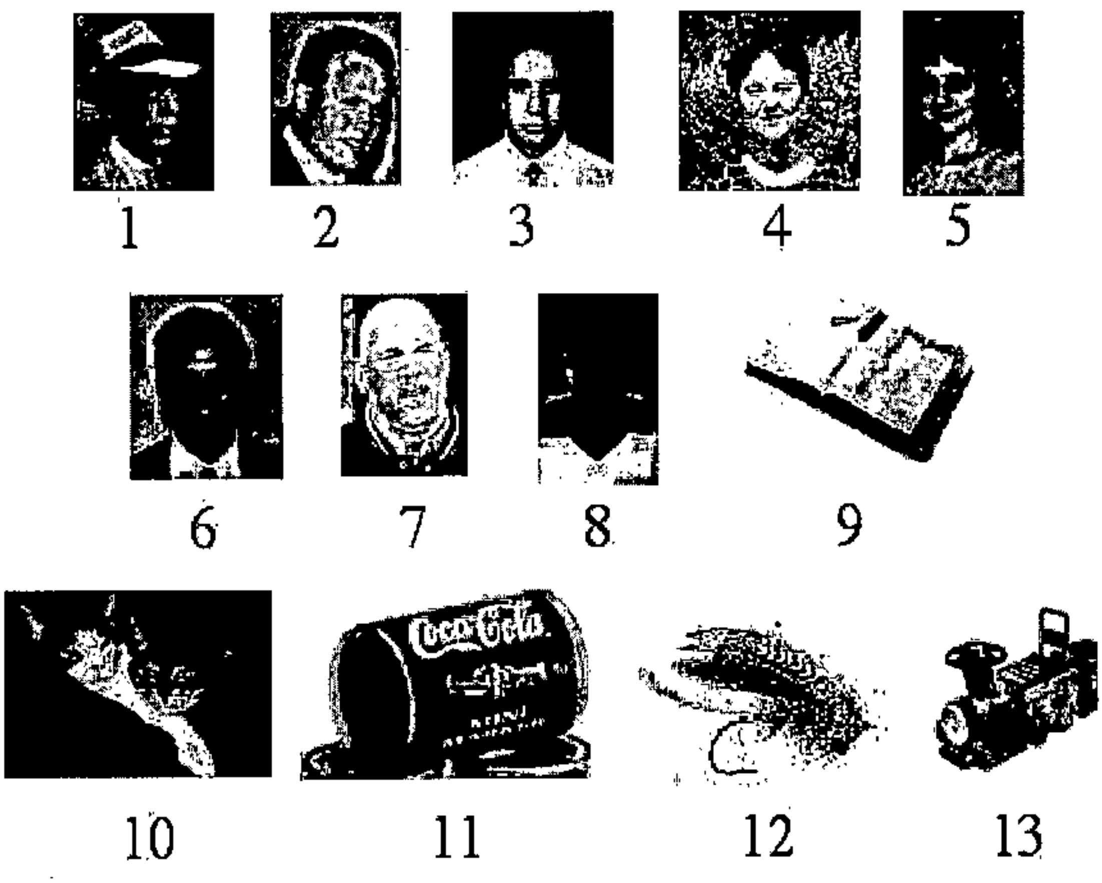
حيث **I** هي المصفوفة الواحدية، $\boldsymbol{\mu}$ معامل التعلّم، **J** هي المصفوفة اليعقوبية لأخطاء الخرج m المتعلقة بالأوزان n للشبكة العصبونية. يتم تعديل المعامل بــ $\boldsymbol{\mu}$ بشكل ذاتي في كل تكرار لضمان التقارب. تتطلب خوارزمية LM حساب J و معكوس المصفوفة المربعة J^TJ(ذات البعد$N\times N$)في كل تكرار.

و للتبسيط ، نحاول فقط أن نكشف عن وجود لقطة لملامح وجه ، وهي وجه بشري من جهة نظر أمامية.بشكل كامل و غير مشوشة في الخلفية. تشمل مجموعة معطيات التدريب على 20 صورة ملونة: 10 وجوه بشرية و10 ليست وجوه بشرية. اختيرت العشر وجوه اختيرت لتمثل تشكيلة من الأعمار و الأجناس و تدرجات ألوان الجلد. أما الصور الأخرى هي عبارة عن أشياء عشوائية مأخوذة من الانترنت. بعض من الصور التي ليست وجوه بشرية اختيرت لخداع الشبكة. على سبيل المثال: الصورة 12 (أقدام)و 15 (يد) و 16 (قفاز بيسبول) تحتوي على تدرجات لون الجلد ، بينما الصورة 11 (كلب) على ملامح وجهية. و سيكون لكل وجه خرج مطابق $P=1$ ، بينما الصور 11– 20 سيكون خرجها $P=0$.



الشكل (2–2) مجموعة صور التدريب

ثانياً، نختار مجموعة من 13 صورة ملونة لتكون مجموعة معطيات الاختبار (الشكل2-3). و نختار عدد مفرد من الصور لنرى فيما إذا كانت النظام سيعمل بشكل صحيح بنسبة فوق أو تحت 50% .إن الشبكة لن تتدرب على هذه الصور، لذلك فإن أداؤها على مجموعة المعطيات هذه سوف يشير إلى فعالية النظام في كشف الوجوه. إن التعريف الصحيح سوف يؤدي إلى القيمة $P > 0.5$ من أجل جميع الوجوه (الصور 1−6) و القيمة $P < 0.5$ من أجل الصور التي ليست وجوه (7−11). مع ملاحظة أن الصور 3 و 6 تحتوي على تدرج للون الوجه غير ممثلة في مجموعة التدريب، لذا فإن هذه الصور ستكون بمثابة تحدي للنظام للتمييز بشكل صحيح.



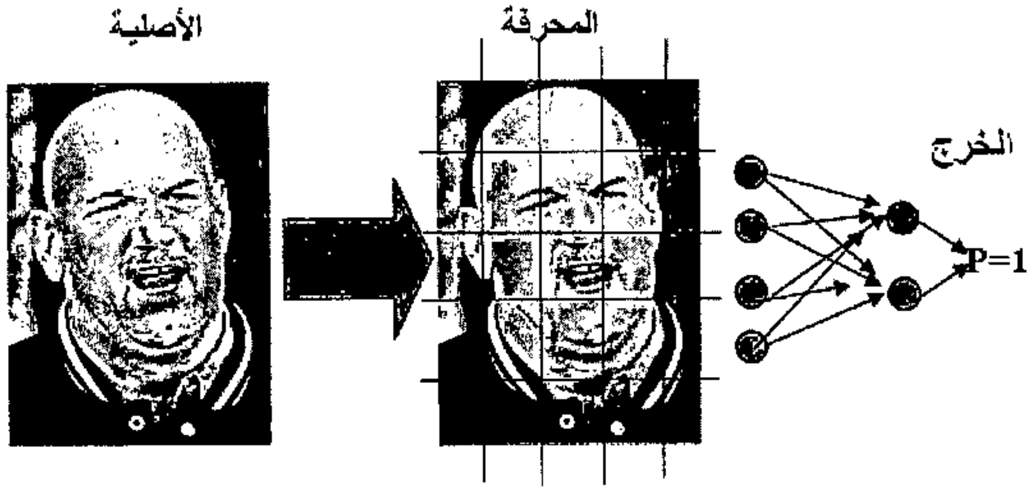
الشكل(2−3) مجموعة صور الاختبار

التفصيل الوحيد المتعلق بالشبكة العصبونية هو تحديد نوع الدخل المخصص لكي تغذى به الشبكة. و ذلك لكون مصفوفة أوزان مجموعة التدريب لها الحجم $20 \times 20(|X|+1) = 400(|X|+1)$. ينبغي لنا إبقاء حجم شعاع الدخل $|X|$ صغير قدر الإمكان للإبقاء على الكلفة الحسابية أقل ما يمكن[19]. و تم اختيار النظام من أجل ثلاث أنماط من أنماط المدخلات:الصور المحرفة و مخططات RGB الإحصائية و مخططات YES

الإحصائية. لذا استعملت مجموعات صور التدريب و الاختبار أعلاه على الأنظمة الثلاث جميعها و ذلك من أجل المقارنة.

3-2 الصور المحرفةInterpolated Images :

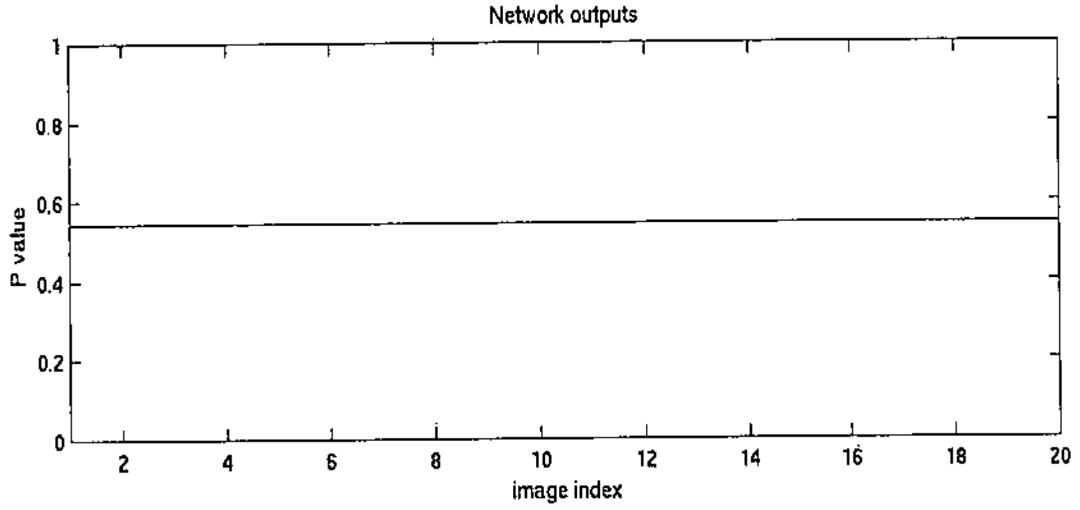
إذا حاولنا القيام بتغذية الشبكة بالصورة الملونة ذاتها كشعاع للدخل X فإن صورة ملونة صغيرة يمكن أن تحتوي على آلاف البكسلات و بالتالي فإنه من غير المنطقي تغذية الشبكة بشعاع دخل كبير X . لذلك يتم تحريف الصورة باختيار شبكة مكونة من N نقطة متباعدة بانتظام على الصورة.و تُضم قيم RGB ضمن شعاع واحد.لذلك فإن شعاع الدخل X سوف يتضمن صفاً، فصف قيم R في شبكة النقط متبوعة بقيم G و أخيراً بقيم B. إذاً حجم شعاع الدخل $|X| = 3N$.



الشكل (2-4) تحريف الصورة من أجل شبكة نقط N=16

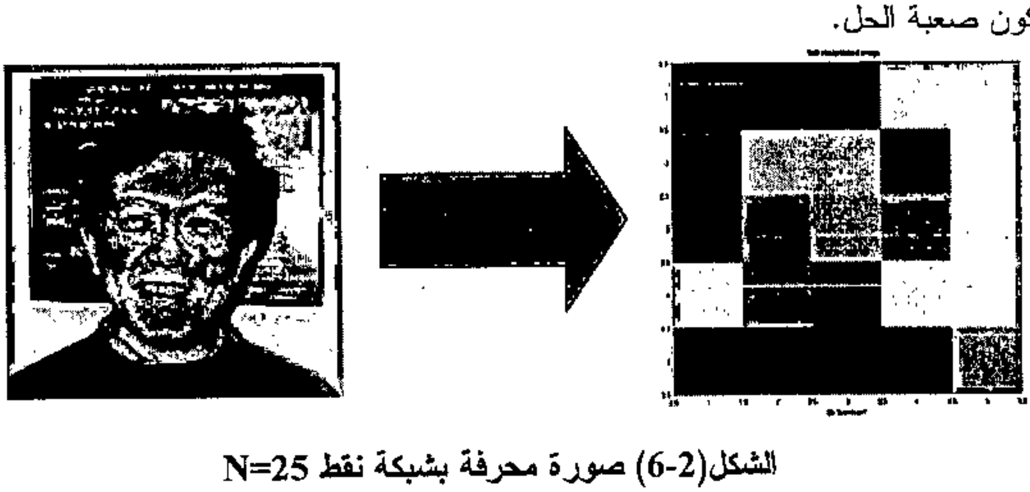
نختار N=25 بشكل تجريبي و بالتالي يكون $|X| = 75$ لكي نحافظ على حجم صغير لشعاع الدخل و بالتالي المحافظة على وقت تدريب منخفض.و ذلك لكون مصفوفة الأوزان لها الحجم $20 \times 20(|X|+1) = 30400$، فإن التدريب سيكون مكلف جداً بشكل حسابي حتى من أجل قيمة N الصغيرة و مجموعة التدريب الصغيرة. وبما أن خوارزمية Levenberg-Marquad تفرض كل قيمة دخل أقل من الواحد و قيمة أي بكسل على الأغلب هي 255 ، فإننا نقسم شعاع الدخل X إلى 260 (و ذلك للتقريب، لتصبح القيمة قريبة من الواحد و لا تبلغه).

مجموعة تدريب الأوزان أدت لأن يصبح الخرج P ثابت. وباستعمال دالة العتبة $\sigma(t) = \frac{1}{1+e^{-t}}$ يجعل من الممكن أن كل دخل سوف يؤدي إلى نفس الخرج .



الشكل(2-5) خرج الشبكة من أجل مجموعة التدريب بعد 100 تكرار لخوارزمية Levenberg-Marquad. الدخل من الصورة المحرفة

و تبعاً لهذا النظام فإن أي دخل سوف يؤدي إلى الخرج $P = 0.5$ بغض النظر إن احتوت الصورة على وجه أم لا.و بما أن الحالة $P = 0.5$تعني أن الخوارزمية لا يمكن أن تحدد احتواء الصورة على وجه. وبالتالي فشل هذا النظام. ربما قد نتج هذا الفشل عن التحريف. كما يمكن أن نرى في الشكل (2-6) أدناه ، صورة محرفة بشبكة نقط N=25 تكون صعبة الحل.

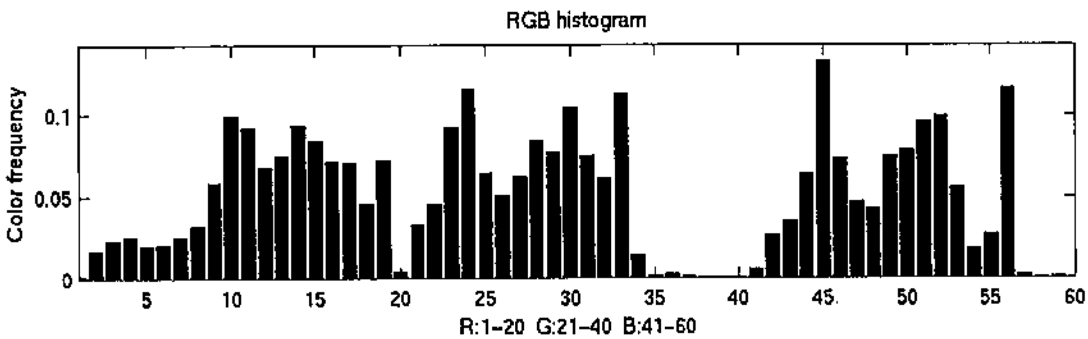


الشكل(2-6) صورة محرفة بشبكة نقط N=25

4-2 مخطط RGB الإحصائي RGB Histogram:

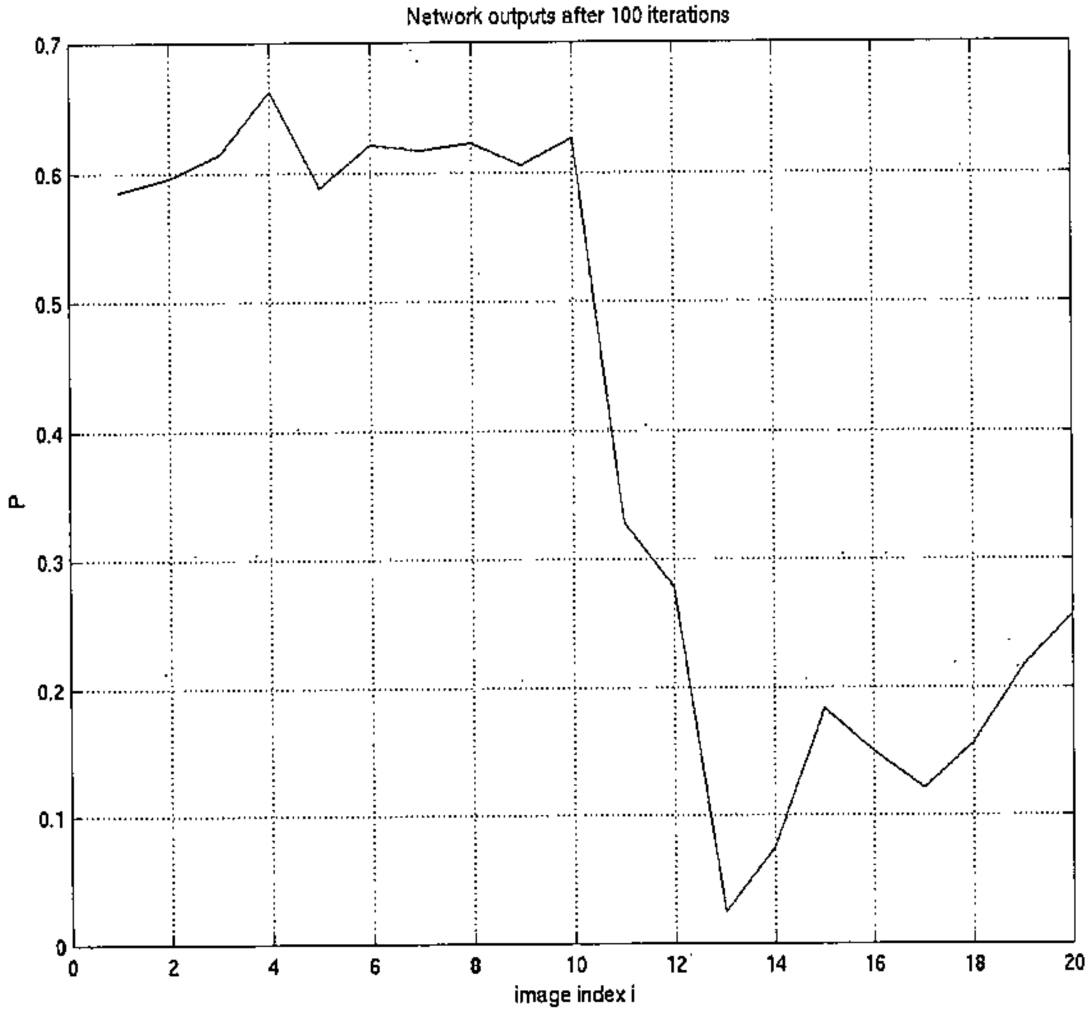
لكون تحريف الصورة لم يعطي النتائج المطلوبة ،سوف ننتقل إلى شعاع الدخل X الذي يمثل الصورة بشكل عام.لقد صنفنا قيم RGB لكل بكسل في الصورة في مخطط التردد النسبي الإحصائي، حيث أن كل مكون من مكونات فضاء اللون(R,G,B) قد مُثل بواسطة N خانة. و مُثلت المخططات الثلاث الإحصائية بعد ذلك كشعاع واحد، لذلك فإن حجم شعاع الدخل سيكون $|X| = 3N$. و كان من المأمول أن تتعلم الشبكة العصبية من الخانات المقابلة لتدرجات لون الجلد .كما نرى في الشكل (2-7)





الشكل (2-7) مخطط RGB الإحصائي حيث N=20

نختار N=20 بشكل تجريبي خانة لكل مكون في فضاء اللون RGB .لذلك فإن شعاع الدخل في هذه الحالة سيكون له الحجم $|X| = 60$.و لقد اتجهت خوارزمية Levenberg-Marquad نحو مجموعة معطيات تدريب من أجل 100 تكرار. و خرج الشبكة نسبة للتدريب كان مناسباً لكون كل صورة مصنفة بشكل صحيح .لذلك فإن الصورة المحتوية على وجوه (1- 10) سوف تعطي الخرج $P > 0.5$ بينما الصور غير المحتوية على وجه سوف تؤدي إلى الخرج $P < 0.5$.



الشكل (2-8) خرج الشبكة من أجل مجموعة التدريب بعد 100 تكرار لخوارزمية Levenberg-Marquad. الدخل مأخوذ بطريقة مخطط RGB الإحصائي.

هذه الطريقة صنفت 7 من 13 من صور الاختبار بشكل صحيح أو 53.8 % .هذا يشير إلى أن النظام أفضل بعض الشيء مما قُدر.على أية حال بقي تحسين هام على نتائج الطريقة الأولى .إن المقياس الشائع لأداء أنظمة التصنيف هو نسب الخطأ السلبي و الخطأ الإيجابي[19]. وعلى الرغم من أن النسبتين من أجل مجموعة التدريب كانتا 0.0 ، فمن أجل مجموعة معطيات الاختبار يكون لدينا:

$$\text{الخطأ الإيجابي} = \frac{\text{عدد الوجوه المكتشفة بشكل خاطئ}}{\text{العدد الكلي للوجوه الفعلية}} = \frac{0}{8} = 0.0$$

$$\text{الخطأ السلبي} = \frac{\text{عدد الوجوه المفقودة}}{\text{العدد الكلي للوجوه الفعلية}} = \frac{6}{8} = 0.75$$

نلاحظ أن تلك الأخطاء كانت مبدئياً من الوجوه المفقودة.هذا قد يكون بسبب أن التدريب يؤدي لأن يصبح خرج الشبكة قريب جداً من 0.5. على الرغم من أن النظام ميّز الوجوه العشر بشكل صحيح في مجموعة التدريب، فإن كل قيم P الموافقة للصور من 1 حتى 10 كانت أكبر بقليل فقط من 0.5 (الشكل 2–8) .ربما نتج هذا عن وجود الصور 12 و 15 في مجموعة التدريب .هذه الصور ليست وجوه لكنها تحتوي على تدرجات لون الجلد. على العموم بأخذ معطيات التدريب و الاختبار معاً، فإن هذه الطريقة ميّزت 27 صورة من 33 بشكل صحيح أي (81.8 %).

2–5 مخطط YES الإحصائي YES Histogram:

على الرغم من إعطاء مخطط RGB الإحصائي أعطى بعض النتائج الإيجابية و لكنه لا يزال يعتبر نظام كشف وجوه غير مقنع .إنّ RGB هو فضاء اللون الطبيعي الذي يتم العمل فيه و ترمّز معظم الصور الملونة بواسطة هذا النظام . و نقوم بتطبيق طريقة مدرج اللون الإحصائي على فضاء اللون YES. فتعطى التحويلات من فضاء RGB القياسي إلى فضاء YES بالمعادلات التالية:

$$Y = 0.253R + 0.684G + 0.063B$$
$$E = 0.5R - 0.5G$$
$$S = 0.25R + 0.25G - 0.5B$$

حيث تختار المصفوفة Y حواف الصورة بينما تقوم المصفوفتان E و S بترميز كثافة اللون.و قد يزود مخطط Y الإحصائي بشكل ما الشبكة العصبية بالمعلومات المكانية بدلاً من أن يزودها فقط بكثافة اللون . لذلك فإن الأخطاء في مخطط RGB الإحصائي ربما تكون متوقعة بشكل جزئي بسبب التشابه بين المخططات الثلاث الإحصائية R و G و B. و تبدو طريقة YES بأنها أدت مخططات إحصائية ذات اختلاف أعظم (الشكل 2–9). وتبعاً لهذه

الطريقة فإن إلحاق المخططات الإحصائية الثلاث كشعاع يزوّد الشبكة بمزيد من المعلومات مثل الكثافة و التباين وفقاً لطريقة YES .

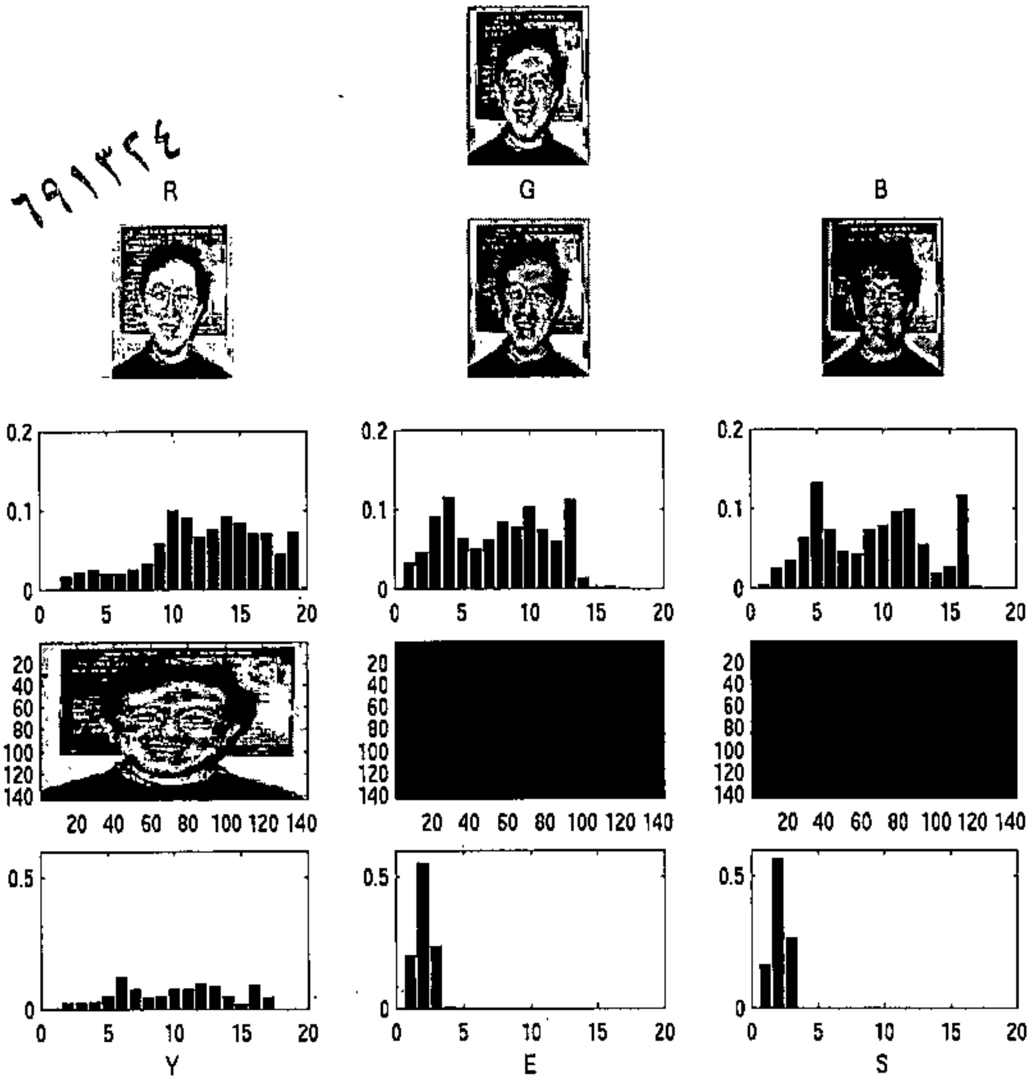
نصنف قيم بكسلات الصورة في مخططات تردد YES النسبية الإحصائية ، كل منها بـــ N خانة موسعة .و يتم إلحاق الإحصائية لتشكيل شعاع دخل واحد X.و نختار N=20 مرة أخرى.

تم تنفيذ خوارزمية Levenberg-Marquad لأجل 100 تكرار بمدخلات من أجل 20 صورة تدريب محددة بطريقة مخطط YES الإحصائي .بعد 100 تكرار ، فإن النظام صنف بشكل صحيح 17 صورة من 20 صورة تدريب الشكل (2−10).صورة واحدة فقط تحتوي على وجه (الصورة 17) كانت قيمة P لها أقل من 0.5 .الصورتان اللتان لم تحتويا على وجه (الصور 15 ، 16) كانت قيمة P لها $0.5 < P < 0.6$.ربما كان سوء التصنيف هذا بسبب كون الصور 15 و 16 تحتوي على تدرجات لون الجلد ، بينما الصورة 5 كانت أغمق درجة لون من تدرجات لون الجلد في مجموعة التدريب. أن الشبكة العصبية قادرة على تصنيف مجموعة صور التدريب بشكل صحيح فيما لو كان التدريب قد نفّذ من أجل عدة تكرارات و لكن ذلك لا يوثر على النتيجة بشكل كبير و يؤدي إلى زيادة الكلفة .و تكون نسب الكشف من أجل مجموعة التدريب:

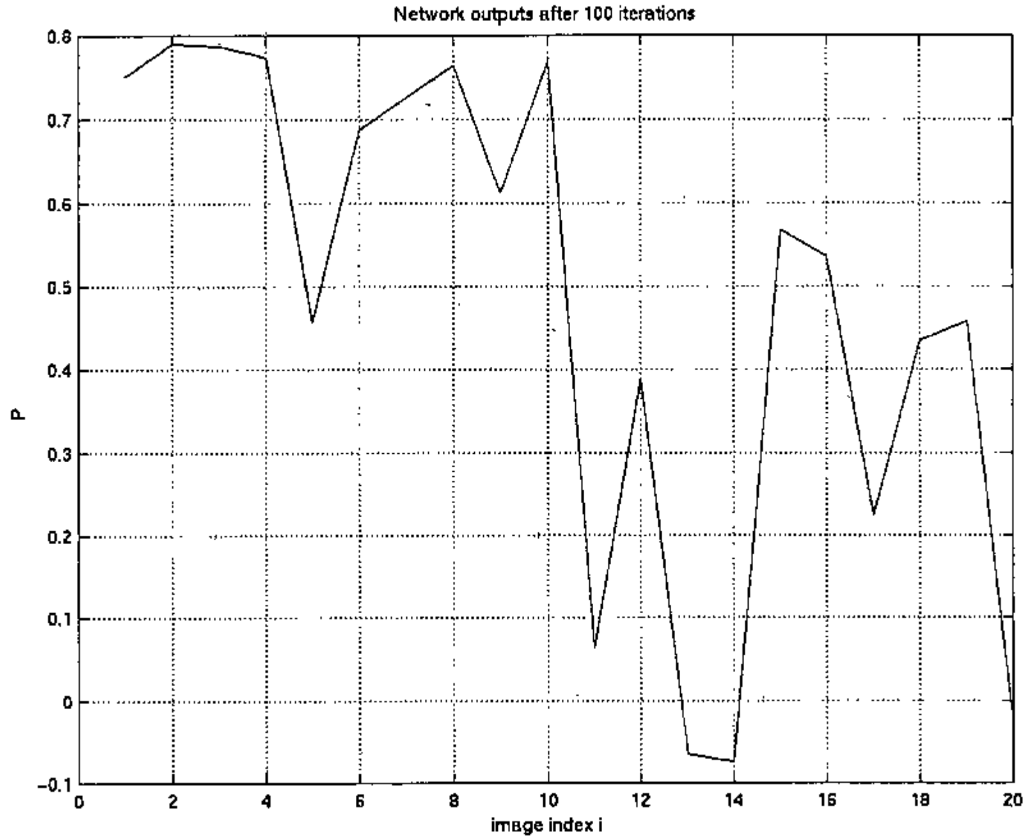
$$\text{الخطأ الإيجابي} = \frac{\text{عدد الوجوه المكتشفة بشكل خاطئ}}{\text{العدد الكلي للوجوه الفعلية}} = \frac{2}{10} = 0.2$$

$$\text{الخطأ السلبي} = \frac{\text{عدد الوجوه المفقودة}}{\text{العدد الكلي للوجوه الفعلية}} = \frac{1}{10} = 0.1$$

إن النتائج من أجل مجموعة معطيات الاختبار كانت أكثر تشجيعاً حيث أن النظام صنف 13 (100 %) من صور الاختبار بشكل صحيح.بشكل عام فإن النظام صنف 30 من 33 (90.9%) من الصور.هذا يعتبر تحسين هام على طريقة مخطط RGB الإحصائي.



الشكل(2-9) التشفير في فضاءات اللون RGB و YES .الصورة الأصلية موجودة في أعلى الشكل.السطر 2 يظهر صور RGB و السطر 3 يظهر مخططات RGB الإحصائية.السطر 4 صور YES و السطر 5 يظهر مخططات YES الإحصائية .



الشكل (2–10) خرج الشبكة من أجل مجموعة التدريب بعد 100 تكرار لخوارزمية Levenberg-Marquad. الدخل مأخوذ بطريقة مخطط YES الإحصائي.

6-2 الخلاصة: Conclusions :

بمقارنة للطرق الثلاث، نجد أن طريقة مخطط YES الإحصائي كانت الأكثر نجاحاً في التمييز الصحيح للوجوه البشرية حيث صنف (90.9%) من الصور. و يبدو لنا أن تشفير الصورة بطريقة YES ينتج معلومات بأعلى جودة بالنسبة للشبكات العصبية.فإنه يُمكن استعمال الطريقة الأولى لاستيفاء الصورة، مما يؤدي إلى أداء أفضل فيما لو ازداد عدد نقاط الشبكة N بشكل ملحوظ. و كل هذه النتائج معتمدة على التدريب البسيط و مجموعة معطيات الاختبار . و نحتاج إلى اختبار الأنظمة على مجموعة معطيات أكبر من أجل حكم أفضل على فعالية تلك الطرق.

إن زيادة حجم N سوف يزود الشبكة العصبية بمعلومات أكثر. زيادة عدد العقد المخفية ربما قد يؤدي إلى تحسين الأداء. مجموعة التدريب الحالية من 20 صورة هي بالأحرى صغيرة.

إن زيادة حجم مجموعة التدريب لتمثيل وجوه مختلفة محتملة يسمح لخوارزمية التدريب بوضع الأوزان بدقة أكبر. وعلى كل حال فإن كل هذه التغييرات سوف تزيد كلف التدريب.

إنه من الممكن أيضاً دمج طريقة مخطط YESالإحصائي وهي طريقة معتمدة على اللون مع طريقة معتمدة على المحتوى .إن الفائدة الرئيسية من طريقة YES هي أنه بعد اكتمال التدريب سيكون النظام سريع جداً حسابياً . و ربما يؤدي دمج هذه التقنية السريعة مع تقنية معتمدة على المحتوى بطيئة لكنها دقيقة إلى نظام كشف وجوه مرغوب.

2-7 النتائج Results و النظرة المستقبلية:

يوضح الجدول(2-1) خرج الشبكة للطرق الثلاث من أجل مجموعة التدريب، حيث أن الطريقة الأولى أعطت نفس القيمة من أجل جميع الصور، و الثانية صنفت100% من صور التدريب بشكل صحيح، أما الثالثة صنفت 85% من صور التدريب.

الجدول(2-1) خرج الشبكة للطرق الثلاث من أجل مجموعة التدريب

دليل الصورة	P الفعلي	الطريقة الأولى الخرج P	الطريقة الثانية الخرج P	الطريقة الثالثة الخرج P
1	1	0.5031	0.5857	0.7513
2	1	0.5031	0.5956	0.7907
3	1	0.5031	0.6138	0.7874
4	1	0.5031	0.6624	0.7746
5	1	0.5031	0.5878	0.4573
6	1	0.5031	0.6211	0.6886
7	1	0.5031	0.6167	0.7260
8	1	0.5031	0.6223	0.7649
9	1	0.5031	0.6051	0.6116
10	1	0.5031	0.6223	0.7690
11	0	0.5031	0.3287	0.0631
12	0	0.5031	0.2784	0.3896
13	0	0.5031	0.0250	0.0650
14	0	0.5031	0.0755	0.0742
15	0	0.5031	0.1831	0.5673
16	0	0.5031	0.1502	0.5351
17	0	0.5031	0.1216	0.2241
18	0	0.5031	0.1564	0.4343
19	0	0.5031	0.2172	0.4577
20	0	0.5031	0.2570	0.0259

الجدول(2-2) خرج الشبكة للطرق الثلاث من أجل مجموعة الاختبار

دليل الصورة	P الفعلي	الطريقة الأولى الخرج P	الطريقة الثانية الخرج P	الطريقة الثالثة الخرج P
1	1	0.5031	0.4769	0.7022
2	1	0.5031	0.3790	0.6704
3	1	0.5031	0.4897	0.7330
4	1	0.5031	0.4420	0.7113
5	1	0.5031	0.4032	0.6934
6	1	0.5031	0.2225	0.5354
7	1	0.5031	0.5597	0.7889
8	1	0.5031	0.5516	0.5778
9	0	0.5031	0.1033	0.1458
10	0	0.5031	0.4851	0.1147
11	0	0.5031	0.2075	0.4777
12	0	0.5031	0.4827	0.3017
13	0	0.5031	0.0066	0.0312

و يوضح الجدول(2-2) خرج الشبكة للطرق الثلاث من أجل مجموعة الاختبار، حيث أن الطريقة الأولى أعطت نفس القيمة من أجل جميع الصور أيضاً، و الثانية صنفت53% من صور الاختبار بشكل صحيح، أما الثالثة صنفت 100% من صور الاختبار.

أما عن النظرة المستقبلية فيمكن أن يطور هذا البحث بحيث يصبح قادراً على تحديد العرق البشري بأقل نسبة خطأ، من خلال تأمين كمية كبيرة من عينات التدريب المحققة للشروط بالإضافة إلى إمكانية تحديد الجنس و العمر لهذا الوجه. بالإضافة إلى أهميته في المختبرات الجنائية.

الفصل الثالث

ضغط الصور الرقمية باستخدام الشبكات العصبونية

1-3 مقدمة Introduction :

يعتبر ضغط الصور من أهم المواضيع في العديد من المجالات مثل الاتصالات و تخزين المعطيات. و قد تم تطوير بعض التقنيات الحسابية الجديدة مثل الشبكات العصبونية لضغط الصور لأنها أثبتت فاعليتها في العديد من المجالات مثل التعلّم و إزالة التشويش و غيرها. لقد كان هناك العديد من المحاولات المقبولة لضغط الصور باستخدام الشبكات العصبونية[20] . العديد من أنماط الشبكات العصبونية جرى استخدامها في أبحاث ضغط الصور مثل Perceptron متعددة الطبقة (Multi-Layer Perceptron) أو MLP [21-31] و هوبفيلد (Hopfield) [32] و خريطة التنظيم الذاتية (Self-Organization Map) أو SOM و تعلّم شعاع القيم المتقطعة (Learning Vector Quantization) أو LVQ[23 -24] و تحليل المكون الرئيسي (Principal Component Analysis) أو PCA[25]. إن شبكة MLP من بين تلك الطرق و التي تستخدم خوارزمية تدريب الانتشار الخلفي تُزود ببنى بسيطة و فعالة.

تم تحقيق ضغط الصور بواسطة شبكة الانتشار الخلفي العصبونية (BPNN)Back Propagation Neural Network في العديد من الأبحاث. و من بينها ما قام به N.Sonehara و M.Kawato وS.Miyake و K.Nakane حيث اقترحوا شبكة من النوع BPNN بثلاث طبقات لضغط الصور[22]. حيث تُقسم الصورة الأصلية في طريقتهم إلى كتل و يتم تغذية عصبونات الدخل بها ، الكتل المضغوطة وُجدت كخرج للطبقة المخفية و الكتل مفكوكة الضغط وُجدت في عصبونات طبقة الخرج. النتائج أثبت أن الشبكة بالطريقة السابقة أعطت نوعية صور سيئة حتى من أجل صور التدريب ونسبة ضغط 4:1[22]. و حاول العديد من الباحثين تطوير أداء الشبكة العصبونية المخصصة لتقنية ضغط الصور [29]. و كانت الشبكات العصبونية الهرمية أحد هذه الجهود التي وسعت BPNN بإضافة طبقتين مخفيتين أو أكثر إليها. و استغل هذا التوسّع الارتباط بين الكتل في الصورة و كذلك الارتباط بين البكسلات في الكتلة الواحدة. و نتج عن هذه الطريقة تحسّن على نسبة SNR للصورة المعاد تركيبها فكان هذا التحسن طفيفاً[33].

تستخدم الطرق التكيّفيّة وجهة نظر أخرى لضغط/فك ضغط (CODEC) كتل الصورة. و يُستعمل لوجهة النظر هذه العديد من الشبكات العصبونية لضغط/فك ضغط كتل الصورة المختلفة مع الأخذ بعين الاعتبار تعقيد الكتل. و اقترح S. Carrato و G. Ramponi و S. Marsi تجميع كتل الصورة في صفوف تبعاً لمقياس التعقيد المسمى فعالية(activity) [24 -25]. و لقد استخدموا أربع شبكات من النوع BPNN مع نسب ضغط مختلفة لكل صف. و قد أدى ذلك إلى تحسين هام على شبكة BPNN الأساسية. و تم اقتراح نظرة تكيّفيّة أخرى تعتمد على استعمال مقياس التعقيد و ذلك بتوجيه الكتلة بواسطة ست شبكات من النوع BPNN و قد أعطت نوعية بصرية أفضل[29]. و بعد استعمال شبكات BPNN في ضغط الصور يتم طرح قيمة كل بكسل من القيمة المتوسطة للكتلة[26]. و تعطي هذه الطريقة بعض النتائج الجيدة و لكنها تتطلب كلف إضافية لإرسال القيم المتوسطة.

سنستخدم في هذا الفصل الشبكة العصبونية الأساسية المعتمدة على خوارزمية لضغط الصور. و من ثم سنقدم نظرة تكيّفيّة لضغط الصور. و سنستخدم طرق لحساب التعقيد في كتل الصورة تعتمد على مفهوم الإنتروبية و الفعالية و مسار الأنماط في الكتل. و تستعمل هذه النظرة التكيّفيّة شبكات مختلفة من النوع BPNN مع نسب ضغط مختلفة و التي تستخدم لضغط/فك ضغط كتل الصورة بالاعتماد على مستوى التعقيد في الكتلة. و لقد استخدمنا بشكل عملي معيار التعقيد لاختيار الشبكة الملائمة لضغط كتلة الصورة القادمة. و تستعمل طريقة أفضل SNR لاختيار الشبكة التي تعطي أفضل SNR لكتلة الصورة تلك. و يستخدم تداخل كتل الصورة بالإضافة لذلك لإزالة تأثير التشويش المسمى(رقعة الشطرنج) في فك ضغط الصورة.و أظهرت نتائجنا التجريبية بنية تداخل الكتل و اختيار الشبكة بأفضل SNR و التي تؤدي إلى تحسين نسبة PSNR. و النوعية البصرية للصورة المعاد تركيبها بالمقارنة مع الترميز JPEG القياسي. و قد تم استخدام صيغتين من صيغ لصور و هي BMP (Bit Map) و JPEG (joint photographic experts group).

سيكون ترتيب هذا الفصل كما يلي: في المقطع الثاني سنناقش شبكة Perceptron العصبونية متعددة الطبقة و نظرتها التكيّفيّة و التي طُورت مباشرة لضغط الصور. كما

سيصف المقطع الثالث طرق قياس التعقيد المستخدمة في هذا الفصل.و سنناقش في المقطع الرابع النتائج التجريبية و أخيراً يتضمن المقطع الخامس ملخص لهذا الفصل.

3-2الشبكات العصبونية المتعددة الطبقة المستخدمة في ضغط الصور

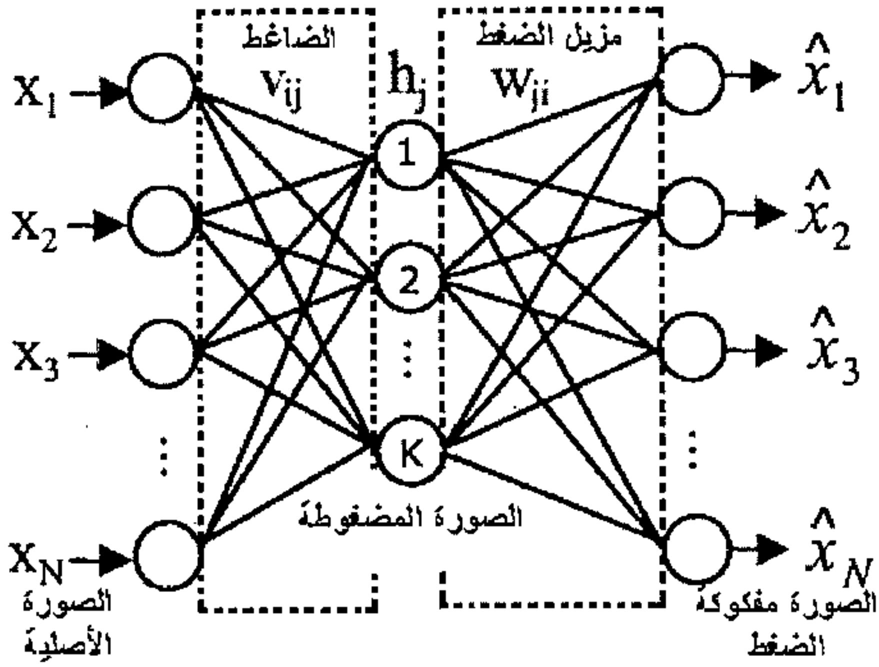
MULTI-LAYER NEURAL NETWORKS FOR IMAGE COMPRESSION

يمكن تطبيق الشبكات العصبونية متعددة الطبقة مع خوارزمية الانتشار الخلفي مباشرة لضغط الصور .إن أبسط شبكة عصبية من أجل هذا الغرض موضحة بالشكل(3-1).و هذه الشبكة مؤلفة من ثلاث طبقات :طبقة الدخل و الطبقة المخفية و طبقة الخرج .ترتبط كل من طبقة الدخل و الخرج بشكل كامل مع الطبقة المخفية و تمتلك كل منهما العدد نفسه من العصبونات N .يُمكن تحقيق الضغط بجعل عدد العصبونات في الطبقة المخفية K أقل من عدد العصبونات في كل من طبقة الدخل و الخرج أي $(K \leq N)$. تُقسم صورة الدخل وفق أفضل طرق الضغط إلى كتل على سبيل المثال إلى 8×8 أو 4 ×4 أو 16× 16 بكسل . يُحدد حجم الكتل عدد العصبونات في طبقات الدخل و الخرج والذي يُحوّل إلى شعاع عمود و تُغذى به طبقة الدخل في الشبكة العصبونية ، عصبون واحد لكل بكسل. فإن الضغط باستخدام شبكة MLP العصبونية الأساسية يتم عبر مراحل التدريب والتطبيق كما يلي:

a) التدريب Training:

يتم في هذه المرحلة مثل جميع عمليات التدريب الأخرى اختيار مجموعة من عينات الصور لتدريب الشبكة باستخدام قاعدة تعليم الانتشار الخلفي. و سيكون النمط الهادف من أجل غرض الضغط في عصبونات طبقة الخرج هو نفس النمط في طبقة الدخل. و يتمثل الضغط باستخدام طبقة مخفية والتي تضغط صورة الدخل عبر ما يسمى بعنق الزجاجة حيث تُغذى طبقة دخل كبيرة طبقة مخفية صغيرة و التي بدورها تغذي طبقة خرج كبيرة مكافئة لطبقة الدخل. و يتم تحويل عينات تدريب الكتل إلى أشعة، و من ثم

يؤخذ نظيمها من مجال تدرجها الرمادي [0,1].تبعاً لتركيب الشبكة العصبونية في الشكل(3-1) .



الشكل(3-1) التركيب الأساسي للشبكة العصبونية المستخدمة في ضغط الصور

يمكن أن نصف عملية تعديل الأوزان من أجل الضغط و فك الضغط بالمعادلات التالية:

$$H_j^{in}=\sum_{i=1}^{N} V_{ij}\ X_i\ ,\ h_j = f(H_j^{in});\ 1\leq J \leq K \qquad (1\text{-}3)$$

$$\hat{X}_i^{in}=\sum_{j=1}^{K} W_{ji}\ h_j\ ,\ \hat{X}_I = g(\hat{X}_i^{in});\ 1\leq i \leq N \qquad (2\text{-}3)$$

حيث f و g هي دوال التفعيل و التي يمكن أن تكون خطية أو غير خطية. V_{ij} و W_{ji} تمثل أوزان الضاغط و مزيل الضغط على الترتيب. و يمكن إيجاد مصفوفة التحويل $N\times K$ المستخرجة من الضاغط والمصفوفة $K\times N$ من مزيل الضغط للشبكة العصبونية الخطية باستخدام تحويل PCA. و يُزود هذا التحويل بالحل الأمثل للشبكة العصبونية من نمط عنق الزجاجة الخطية المستخدمة لضغط الصور و التي تعطي أصغر قيمة لمتوسط مربع الخطأ بين الصورة الأصلية و الصورة المعاد تركيبها. و يتم تمثيل عينات الدخل في فضاء جديد و

الذي فيه كل العينات غير مرتبطة ، و هذه الحقيقة تؤدي إلى أفضل ضغط. و لكن هذا يعتمد على تحويل المعطيات و يمكن أن تزود بأفضل ضغط لصور التدريب فقط. لذا فإن استخدام دوال التنشيط الخطية و غير الخطية في هذه الشبكة سينتج تحويلات PCA الخطية و غير الخطية على الترتيب.

إن عملية تدريب الشبكة العصبونية ذات البنية المبينة بالشكل(3-1) تكرارية و تتوقف عندما تتقارب الأوزان من قيمها الحقيقية. و يتوقف التدريب في التطبيقات الحقيقية عندما يبلغ الخطأ في العلاقة(3-3) العتبة المسماة ε أو عدد أعظمي من التكرار يحدد تكرار عمليات التدريب.

$$Err=\frac{1}{2}\sum_{k=1}^{N}(X_k - \hat{X}_k) \qquad (3\text{-}3)$$

b) التطبيق Application :

يتم تغذية الشبكة بصورة الاختبار والحصول على الصورة المضغوطة من خرج الطبقة المخفية عند انتهاء التدريب و تعديل الأوزان النهائية. و يجب تثبيت قيم نواتج الخرج تلك عند العدد المطلوب من الخانات.فإذا أُستعمل العدد نفسه من الخانات لتمثيل عصبونات طبقة الدخل و الطبقة المخفية، و بالتالي ستكون نسبة الضغط (Compression Ratio) أو CR ستكون نسبة عدد عصبونات الدخل إلى العصبونات المخفية.على سبيل المثال لضغط كتلة صورة من الحجم 8 X8 يتطلب 64 عصبون في كل من طبقتي الدخل و الخرج. و في هذه الحالة لو كان عدد العصبونات المخفية 16 (بمعنى آخر كتلة الصورة من الحجم 4X 4) و بالتالي ستكون نسبة الضغط 64:16=4:1.في حال استخدام 32 بت المرمزة بنمط الفاصلة العائمة في تشفير الصورة المضغوطة فإن نسبة الضغط ستكون 1:1 و هذا يشير إلى الضغط و كأنه لم يحدث. و بشكل عام فإن نسبة الضغط للشبكة العصبونية الأساسية و الموضحة بالشكل(3-1) لصورة مؤلفة من n كتلة ستكون كما في العلاقة التالية:

$$CR=\frac{nNB_I}{nKB_H}=\frac{NB_I}{KB_H} \qquad (4\text{-}3)$$

حيث B_I و B_H هما عدد البتات اللازمة لتشفير خرج طبقة الدخل و الطبقة المخفية على الترتيب. و في هذه العلاقة N و K هما عدد العصبونات في طبقة الدخل و الطبقة المخفية على الترتيب. و تُحوّل الصورة المضغوطة في مزيل الضغط إلى نسخة مشابهة للصورة الأصلية بتطبيق الأوزان بين الطبقة المخفية و طبقة خرج مزيل الضغط على نواتج الطبقة المخفية. يجب تخفيض نواتج طبقة الخرج إلى مدى تدرج الرمادي الأصلي أي [0~ 255] من أجل 8 بكسلات.

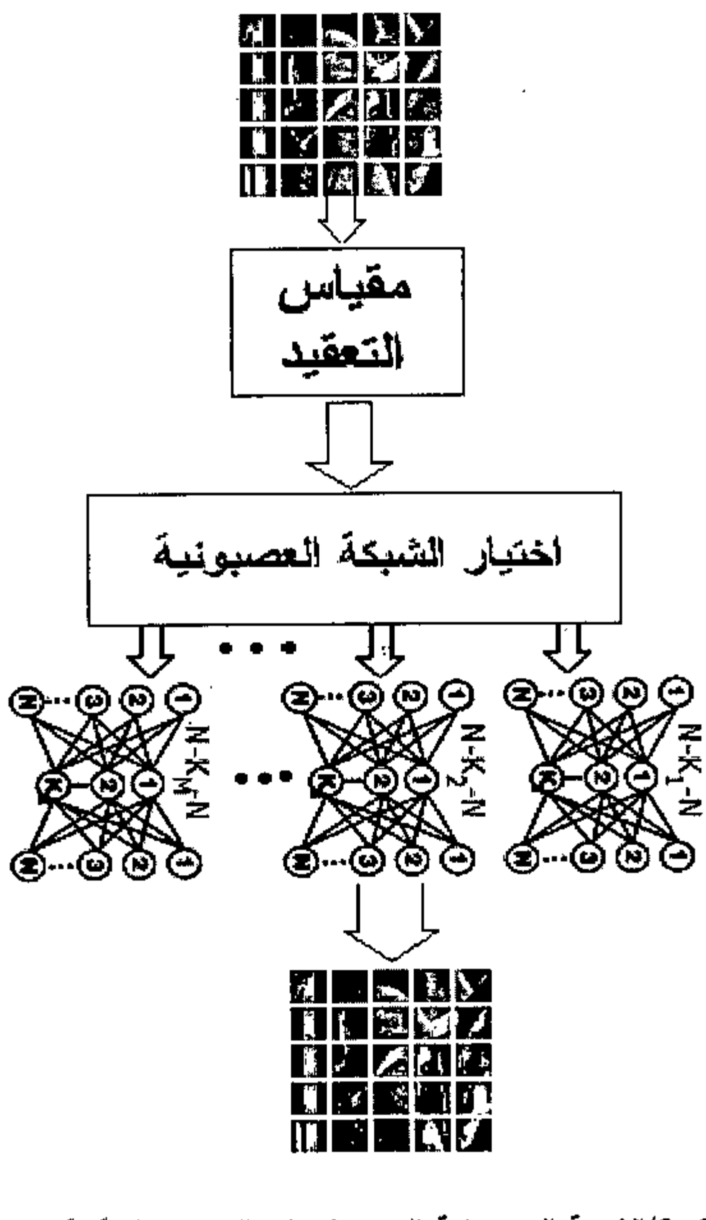
C) النظرة التكيّفية Adaptive approach :

يُزوّد التركيب الأساسي للشبكة العصبونية لضغط الصور كما هو مذكور في المقطع السابق بتقريب لتحويلات PCA.يقوم هذا التركيب بإزالة الترابط بين بكسلات عينات الدخل،هذه العملية أساسية في مجال ضغط البيانات.إن هذا هو السبب الرئيسي لاستبدال PCA بأقرب تقريب لها و هو تحويل جيب تمام (Discrete Cosine Transform) أو DCT للبيانات المستقلة في التطبيقات الحقيقية. و بسبب هذه المحدودية في تركيب الشبكة العصبونية الأساسية المستخدمة للضغط، فإن النتائج التي تم الحصول عليها من هذه الشبكة تظهر بأنها ضعيفة جداً في حال استعمالها في الضغط.

إن النظرة التكيّفية هي واحدة من الطرق المستخدمة لتحسين أداء ذلك التركيب البسيط و التي تستخدم شبكات عصبونية مختلفة لضغط كتل الصورة[21-24-29].و ليتم ذلك أولاً تُقسم الصورة إلى عدة طبقات تبعاً لتعقيدها. و من ثم تُستخدم كتل الصورة لكل طبقة.و أن لكل الشبكات التركيب ذاته، و لكنها تختلف في عدد عصبونات الطبقة المخفية والتي ستعطي نسب ضغط مختلفة.

باعتبار الشبكة العصبونية الموضحة بالشكل(3-1) كتركيب أساسي ، يمكن تقديم الطريقة التكيّفية كما في الشكل(3-2).من الشكل نلاحظ أنه لتدريب الشبكات على المعلومات المتوفرة في كل طبقة يتم تقديرها بقيمة متوسطة تبعاً لمقياس تعقيد مثل تدرجات اللون الرمادي في كتلة الصورة أو بطرق أخرى. ثم تبعاً لقيمة التعقيد تلك يتم اختيار واحدة من الشبكات المتوفرة.تُدرّب كل شبكة باستخدام معطيات التدريب المطابقة بخوارزمية الانتشار المرتد. و للتمييز بالنسبة لمزيل الضغط أي شبكة تم استخدامها في مرحلة الضغط لضغط كتلة الصورة

يخصص ترميز معين لكل شبكة مدربة.و هذا الترميز يجب أن يرسل أو يحفظ على امتداد الصورة المضغوطة.و من الواضح أن عدد الشبكات و بالتالي عدد البتات المطلوب إعطائه لذلك الترميز سيؤثر على نسبة الضغط. و يُفضّل العدد المنخفض من البتات و لكنه من جهة أخرى العدد المنخفض من الشبكات يقلل من تكيّف الخوارزمية. و في مزيل الضغط يتم استخراج الترميز المحفوظ أو المرسل على امتداد الصورة المضغوطة و لذلك فإن الشبكة المطابقة(أي نفس الشبكة المستخدمة في مرحلة الضغط) يمكن أن يتم اختيارها لإزالة الضغط.



(الشكل 3-2)الشبكة العصبونية المعتمدة على التركيب التكيّفي لضغط الصور

نفترض في النظرة التكيّفية وجود M شبكة مختلفة بـــ k_1 - k_M عصبون في الطبقة المخفية. في هذه الحالة من أجل صورة مكونة من n كتلة كل منها مكونة من N بكسل و تكون عندها نسبة الضغط كما في العلاقة(3-5) و التي يتم الحصول عليها بتعديل العلاقة(3-4) .

$$CR_a = \frac{nNB_I}{\sum_{i=1}^{n} B_H K_j^i + q} = \frac{NB_I}{\left(\frac{B_H}{n}\sum_{i=1}^{n} K_j^i\right) + \frac{q}{n}} \;;\; 1 \le j \le M \quad (5\text{-}3)$$

حيث K_j^i هو عدد العصبونات في الطبقة المخفية للشبكة المختارة من أجل كتلة الصورة i حيث $1 \le j \le M$، q عدد البتات المطلوبة لترميز رقم الشبكة حيث q يساوي العدد الصحيح الأصغر بحيث يكون $2^q \ge M$.

3-3 طرق مقياس التعقيد

COMPLEXITY MEASUREMENT METHODS

سنقدم الطرق الثلاث التالية لحساب مستوى التفصيل في كتل الصورة لدمجها في خوارزمية BPNN التكيّفية. و تُصنف كتل الصورة إلى صفوف متعددة بالاعتماد على قيم مستوى التفصيل. كما تُخصص شبكة واحدة لكل صف و نسبة الضغط لذلك الصف ترتبط بتعقيده. و بما أن مقياس التعقيد يجب أن يكشف كمية المعلومات في كتلة الصورة ،فيجب أيضاً أن تميّز كتل الصورة حسب الشبكة العصبونية المستخدمة في الضغط. و إن معيار مقياس التعقيد هو عامل مهم لهذه النظرة، فهو يؤثر على أداء الضغط بشكل ملحوظ .و هنا نستخدم ثلاث معايير مختلفة هي الإنتروبية و المعتمدة على النمط و الفعالية.

a) التعقيد المعتمد على الإنتروبية Complexity based on Entropy:

تُعتبر الإنتروبية معيار هام لقياس كمية المعلومات ضمن مجموعة من الرموز مثل الصورة. و أن إنتروبية كتلة صورة ذات N مستوى مختلف من تدرج الرمادي تحسب بالعلاقة التالية:

$$\text{Entropy} = -\sum_{i=1}^{N} P(x_i)\log P(x_i) \qquad (6\text{-}3)$$

حيث $P(x_i)$ هو احتمال ظهور المستوى x_i من تدرج الرمادي في هذه الكتلة.

تحتوي كتلة الصورة التي لها قيمة إنتروبية مرتفعة على معلومات أكثر. و هذا يعني لمنع فقدان المزيد من المعلومات، يجب أن تمر تلك الكتلة بأقل ضغط ممكن.و أيضاً فإن لكتلة ما قيمة إنتروبية منخفضة يجب أن تضغط بالشبكة التي تزود بنسبة ضغط مرتفعة.

b) التعقيد المعتمد على الفعالية Complexity based on Activity:

تُعرف طريقة المقياس لتغطية الفكرة غير الموضوعية للفعالية في كتلة الصورة. من أجل كتلة صورة مكونة من N بكسل(هذا يعني أنها من الحجم $\sqrt{N} \times \sqrt{N}$) فإن الفعالية تعرف بالعلاقة التالية:

$$\text{Activity} = \sum_{i=2}^{\sqrt{N}} \sum_{j=2}^{\sqrt{N}} \left[\sum_{m=-1}^{1} \sum_{n=-1}^{1} (x_{i,j} - x_{i-m,j-n})^2\right] \quad (7\text{-}3)$$

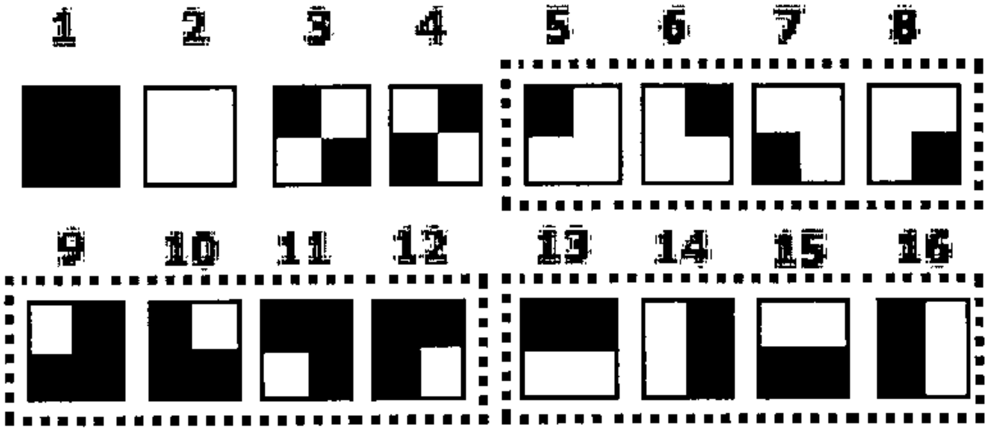
حيث i,j زوجيان ؛ $(m,n) \neq (0,0)$

تتطلب الصفوف ذات قيم الفعالية المنخفضة في هذا المقياس شبكات ذات نسب ضغط مرتفعة و تتطلب الصفوف ذات قيم الفعالية المرتفعة المحافظة على المزيد من المعطيات و بالتالي يجب استعمال شبكات ذات العدد الأكبر من العصبونات المخفية و نسبة الضغط الأخفض.

c) النمط المعتمد على التعقيد Pattern-based complexity :

على الرغم من أن الفعالية معيار جيد في تقريب التعقيد و الإنتروبية هي مقياس دلالي لحساب كمية المعلومات في كتلة الصورة، و لكن استخدامنا لهذه الطرق بغرض التعلّم سيجعلنا نواجه بعض الصعوبات. و أنه من الواضح أن قيمة الإنتروبية لكتل الصورة (من 5 إلى 8) في الشكل(3–3) متساوية. و هذا ينطبق على قيم الفعالية أيضاً. و سبب ذلك أن تلك الكتلِ لديها العدد نفسه من البكسلات البيضاء و السوداء على الرغم من أن كل منها لديه شكل يختلف عن الكتل الأخرى.و بالتالي إذا استُخدم معيار مقياس الإنتروبية لاختيار الشبكة الملائمة فإن كل تلك الكتل سوف تضغط بالشبكة ذاتها. و لدى كل من هذه الكتل نمط مختلف، و لكي نحصل على أفضل نسبة ضغط فإنه من الأفضل تخصيص هذه الكتل بشبكات مختلفة.

و بالتالي نستنتج بأن الإنتروبية و الفعالية لا تستطيعان تمييز هذه الأنماط المختلفة.و أسباب ذلك الاختيار الخاطئ للشبكة الملائمة في الخوارزمية و التقدير السيئ لتقريب التعقيد.



الشكل(3-3) الأنماط المستخدمة في تصنيف كتل الصور وفق معيار مقياس التعقيد

للتغلب على هذه المشكلة نستعمل مقياس تعقيد آخر يسمى بالطريقة المعتمدة على النمط. و تُصنف كتل الصورة بالاعتماد على أنماطها.و يتم ذلك بتقسيم الكتلة إلى أربع كتل جزئية متساوية. و تعتمد طريقة التقسيم هذه على طريقة تمثيل الشجرة الرباعية، و بالتالي تخصص كل كتلة من كتل الصورة بنمط واحد من الأنماط الستة عشر كما في الشكل(3-3).إن الكتل جزئية هي إما بيضاء أو سوداء و من الضروري استخدام عتبة لتحديد مستوى الأبيض أو الأسود لتدرج الكتل الجزئية الرمادي. و تُستخدم شبكة واحدة لكل تلك الأنماط من أجل ضغط كتلة الصورة الموافقة. و للشبكات نسب ضغط مختلفة حسب الأنماط المتعلقة بكل شبكة منها.و بهذا فإن الأنماط 1 و 2 في الشكل(3-3) لها نسب الضغط العظمى، و الأنماط 3 و 4 لها نسب الضغط الصغرى.

3-4 النتائج التجريبية EXPERIMENTAL RESULTS

سنقوم بهذا المقطع بتقييم قابلية الضغط للشبكة العصبونية الأساسية ذات البنية الموضحة بالشكل(3-1) و النظرة التكيّفية المقترحة في الشكل(3-2) بمعيار مقياس تعقيد مختلف. و المقارنة أيضاً بين الطريقة التكيّفية و خوارزمية تشفير JPEG القياسية.بالإضافة إلى نسبة الضغط (CR) المعطاة بالعلاقتين(3-4) و (3-5) فإن أداء تلك الطرق يتم مقارنته حسب معيار النسبة (قمة الإشارة إلى التشويش) أو (Peak Signal to Noise Ratio) و التي يرمز

لها PSNR. يستعمل PSNR غالباً لبساطته في الحساب كمعيار يبين نوعية الصورة المولّدة بواسطة الضغط الذي يفقدها بعض من جودتها. و أن هذه الطريقة بغض النظر عن بساطتها لا تتعلق بشكل محدد بنوعية الصورة المضغوطة.و بالتالي تُفرض الصورة الأصلية X كإشارة نقية والتي تُعتبر صورتها المضغوطة $\hat{X}$ كإشارة مشوشة. و يُفترض أن تكون كلا الصورتين الأصلية و المضغوطة من نفس الحجم.من أجل صورة ذات حجم $R \times C$ (صف × عمود) فإن PSNR تعطى بالعلاقة(8-3) مقدرة بالديسيبل.

$$PSNR = 10\,\log_{10}\left(\frac{255^2}{\frac{1}{R\times C}\sum_{i=1}^{R}\sum_{j=1}^{C}(X_{ij}-\hat{X}_{ij})}\right)\ (dB) \quad (8\text{-}3)$$

لقد استخدمنا في تجاربنا صورة ذات تدرج رمادي ذات دقة 8 بت / بكسل،

حيث تشير 255 في العلاقة (8-3) إلى مستوى الرمادية الأقصى. إن حجم كتل الصورة هو أحد المتغيرات التي من الضروري أن تُحسّن. لقد استخدمنا 4 × 4 و 8 × 8 و 16 × 16 كحجم لكتل الصورة و التي تعطي على الترتيب 16 و 64 و 256 عصبون في كل من طبقتي الدخل و الخرج. و قد تم تقييم حجم الكتلة بالقيم السابقة بالتركيب الأساسي الموضح بالشكل(3-1). و إجمالاً فإن هذه الشبكات تثبت نسبة الضغط 4:1 (أي 16:4 و64:16 و 256:64 على الترتيب) و التي تقابل 4 و 16 و 64 عصبون في الطبقة المخفية (أي كتل من الأحجام 2 × 2 و 4 × 4 و 8 × 8). و تظهر النتائج قيمة هذا المتغير ليست حاسمة، و لكن هناك بعض الاعتبارات حوله. و ينتج الحجم الأكبر للكتلة العدد الأكبر من المتغيرات و تتطلب المزيد من أنماط التدريب. و هذا يؤدي أيضاً إلى أعلى تباين للخطأ ونسبياً أفضل PSNR من أجل ضغط الصور التي تنتج عن مجموعة التدريب. و من الواضح أن حدث عكس ذلك من أجل أحجام الكتلة الصغيرة. و أن السرعة هي المتغير الآخر الذي يبدو أنه يتأثر بحجم الكتلة. و أن عدد الكتل في صورة ما ينقص بازدياد حجم الكتلة، لذا يبدو أن سرعة الخوارزمية تتحسن عندما يكون حجم الكتلة أكبر.و لكن الاعتبارات التالية تنقض الكلام السابق.

بفرض صورة مكونة من N بكسل مقسمة إلى k_s كتلة، كل كتلة من الحجم $\sqrt{N_s} \times \sqrt{N_s}$. و يوجد في تركيب الضغط الأساسي عندما تكون CR هي (4:1) K_s مرة بحيث

يكون شعاع من الحجم $N_s \times 1$ مضروباً بمصفوفة أوزان الضاغط من الحجم $\frac{N_s}{2} \times N_s$. و مع الأخذ بعين الاعتبار أيضاً جداء الكتل المضغوطة ذات الحجم $\sqrt{\frac{N_s}{2}} \times \sqrt{\frac{N_s}{2}}$ بمصفوفة أوزان مزيل الضغط ذات الحجم $N_s \times \frac{N_s}{2}$. و يكون عدد الجداءات و المجاميع معطى بالعلاقتين (3-9) و (3-10) على التوالي.

$$R_S^m = k_s\,[(N_s \,.\, \frac{N_s}{2}) + (\frac{N_s}{2} \,.\, N_s)] = N.N_s \qquad (9\text{-}3)$$

$$R_S^s = k_s\,[((N_s - 1) \,.\, \frac{N_s}{2}) + ((\frac{N_s}{2} - 1).\, N_s)] = N.(N_s - \frac{3}{2}) \qquad (10\text{-}3)$$

حيث $k_s = \frac{N}{N_s}$. إذا تم استخدام حجم كتلة أكبر $\sqrt{N_l} \times \sqrt{N_l}$ حيث $N_l = N_s$ و بالتالي العلاقتين (3-9) و (3-10) تصبحان كما يلي:

$$R_l^m = N.(2.N_s) = 2.R_S^m \qquad (11\text{-}3)$$

$$R_l^s = N.(2.N_s - \frac{3}{2}) = 2.N.(N_s - \frac{3}{4}) \approx 2.R_S^s \qquad (12\text{-}3)$$

هذا يعني أنه لا يمكن الحصول على تحسين للسرعة باستخدام حجم كتلة أكبر حيث كما نلاحظ من العلاقتين السابقتين تضاعف عدد الجداءات و المجاميع. و بشكل عام فإن حجم كتلة الصورة ليس متغير معياري و من بين أحجام الكتلة الثلاث المجربة استخدمنا الحجم8 ×8 في جميع تجاربنا.

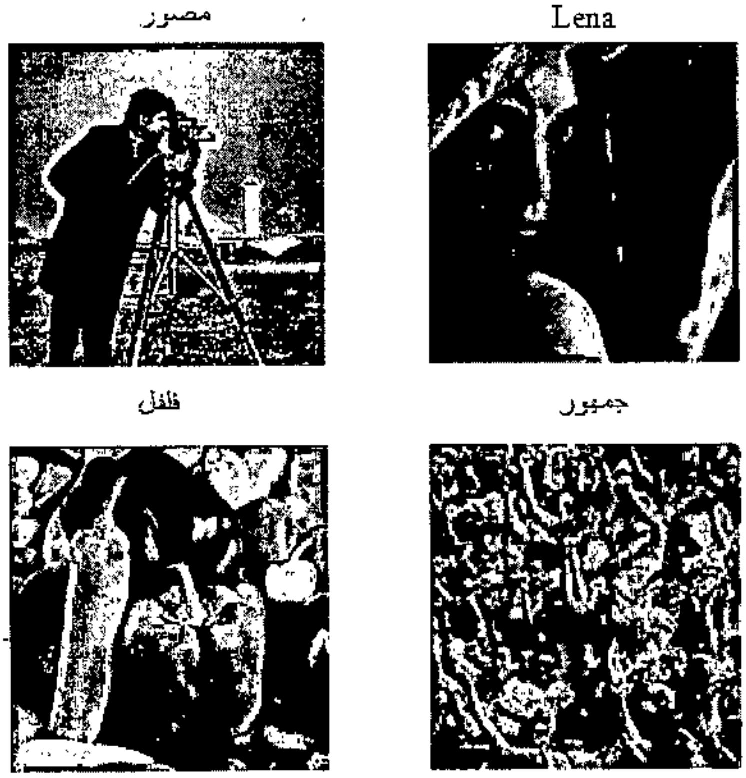
a) تقييمات تركيب الشبكة الأساسي

Evaluations on basic network structure

نستخدم النسبة CR 4:1 لتقييم تركيب الشبكة الأساسي.كتل صورة الدخل من الحجم 8 ×8 و كل بكسل مكون من 8 بت و نفترض 8 بت/ عصبون لترميز خرج الطبقة المخفية. و نستخدم صورة Lena من الحجم 256×256 لتدريب الشبكة بدون استخدام التراكب بين كتل الصورة. و تُستخدم صورة Lena مع ثلاث صور أخرى كصور اختبار. و يوضح الشكل(3-5) نتيجة CODEC باستخدام التركيب الموضح بالشكل(3-1) و يبين الجدول(3-1) النسبة PSNR و التي يتم الحصول عليها بمقارنة الصور الأصلية مع نسختها المضغوطة.لذا فإن نتائج BPNN أدائها أفضل بالنسبة لصورة التدريب بالمقارنة مع الصور الأخرى في

مجموعة الاختبار. و أن هذه الطريقة لا تُزود بجودة مقبولة بالنسبة لصور الاختبار على عكس مجموعة التدريب. و أن التركيب الأساسي يقوم بحساب تحويل PCA للصورة المُدرّبة فقط.

هناك بعض الاعتبارات التي يجب مراعاتها عند تطبيق التركيب الأساسي و الذي يمكن أن يؤثر على النتائج. فإن تهيئة متغيرات الشبكة و عدد البتات في الطبقة المخفية و عدد مرات تدريب المعطيات من أهم المفاهيم من أجل المزيد من الاستكشاف. و أن الأوزان الأولية لتركيب الشبكة العصبونية تعتبر نقطة البداية لإيجاد التحويل الأمثل. و لقد استخدمنا التهيئة العشوائية في تجاربنا، و لكن هناك أعمال أخرى مثل " ضغط الصور باستخدام نموذج الشبكة العصبونية " و الذي يختبر قيم أخرى[22]. و لقد تم استخدام شبكة عصبونية ذات تغذية أمامية بثلاث طبقات للضغط و إزالة الضغط و قيم DCT المجرّبة كأوزان أولية.لقد أظهر تقييمها أن هذه التهيئة لم تعطي أي تحسين على نوعية الصورة المفكوكة الضغط.يؤثر عدد البتات في الطبقة المخفية على نسبة الضغط و نوعية الصورة المعاد تركيبها. و قد استخدمنا 8 بت في خرج الطبقة المخفية[22]. و في بحث " ضغط الصور باستخدام نموذج الشبكة العصبونية " كان عدد البتات في الطبقة المخفية بين 1 و 10 و في بحث " الشبكات العصبونية الاصطناعية لضغط الصور " كان هذا العدد بين 1 و 8[23]. فقد تم الوصول إلى نفس النتيجة تقريباً في كلا البحثين و الأداء لم يزداد بشكل ملحوظ باستخدام أكثر من 6 بت و لكن ذلك حسّن نسبة الضغط. على الرغم من عدم وجود قاعدة معينة لكمية بيانات التدريب بالنسبة للشبكة العصبونية فإن ذلك يتطلب من 4 حتى 10 من عينات التدريب. و لقد أخذنا ذلك بعين الاعتبار في تجاربنا. و بشكل عام فإن الزيادة في حجم مجموعة التدريب سيؤدي إلى الزيادة في خطأ التعلّم و لكن يُنقص خطأ الاختبار. فبزيادة عينات التدريب تتقارب قيم كل من الخطأين السابقين إلى نفس القيمة.



الشكل(3-4) مجموعة صور الاختبار(Lena أيضاً موجودة في مجموعة التدريب)



الشكل(3-5)مجموعة صور الاختبار المعاد تركيبها باستخدام الشبكة الموضحة بالشكل(3-1)

b)نتائج الطرق التكيّفيّة **Results of adaptive methods:**

لقد استخدمنا النظرة التكيّفيّة الموصوفة في الفقرة (2) الجزءc .تحتاج هذه الطريقة إلى المزيد من معطيات التدريب بسبب العدد الأكبر من متغيرات الشبكة.يتم اختيار مجموعة التدريب بالطريقة التي تغطي تشكيلة واسعة من التعقيدات في الصور.و أيضاً عدد أنماط التدريب لجميع الشبكات يجب أن يكون كافي للوصول إلى التقارب بالحد الأدنى للخطأ. و لقد اخترنا 105 صورة و التي تحتوي على تشكيلة واسعة من التعقيدات من أجل جميع مستويات الرمادية.و يظهر الشكل(3-6) بعض من عينات صور التدريب التي اخترناها.

الجدول(3-1) نتائج الضغط باستخدام تركيبEBPNN الأساسي

صورة الاختبار	PSNR(dB)
Lena	34.92
مصور	26.67
جمهور	23.41
فلفل	22.20

تم استخدام كتل صورة من الحجم 8 ×8 في تدريب الشبكات التكيّفيّة كما في التركيب الأساسي، أي 64 عصبون في طبقات الدخل. تُعطي كل كتلة شبكتها الموافقة تبعاً لكل طريقة من طرق قياس التعقيد. و تُعطي الكتل ذات المستوى المنخفض من التعقيد شبكات بأعلى نسبة ضغط و الكتل ذات المستوى العالي من التعقيد تعطي شبكات بأدنى نسبة ضغط. و يُحدد عدد العصبونات في الطبقة المخفية في هذه التراكيب نسبة الضغط و تتعلق بقيمة التعقيد في الكتل. و لقد اخترنا هذه القيمة من أجل كل شبكة في كل معيار من معايير التعقيد كما هو مناقش فيما بعد.

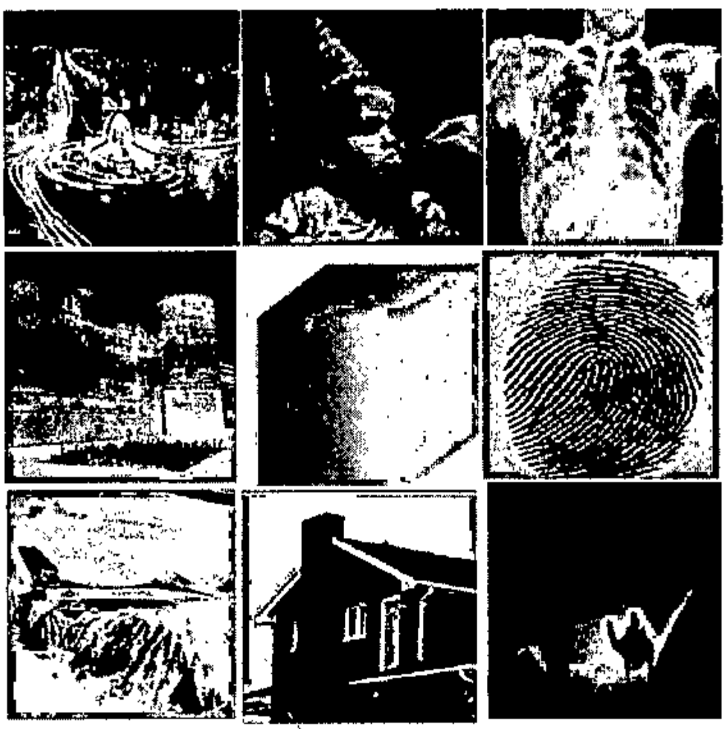
يتم إنجاز الضغط كتدريب اعتيادي و ذلك بأخذ الشبكات المدرّبة و تغذيتها بصورة اختبار. و تُقسّم الصور إلى كتل و يتم حساب قيمة التعقيد لكل كتلة.و بعد ذلك فإن الشبكة المتعلّقة بقيمة التعقيد تلك يتم اختيارها و يتم إنجاز الضغط بواسطة تلك الشبكة. و لقد دعونا هذه الطريقة في تجاربنا بالمعتمدة على التعقيد.

إن اختيار الشبكة المثلى لضغط كتلة الصورة من بين الشبكات المدرّبة هو الخيار الآخر. و لا يُراعى مقياس التعقيد في هذه النظرة و الشبكة التي يتم اختيارها في هذه الحالة

تقلل من معيار مقياس الخطأ. و لقد استخدمنا طريقة اختيار الشبكة هذه مع الأخذ بعين الاعتبار جعل المعيار SNR في حده الأعلى من أجل كل كتلة. و يتم هنا يتم إعطاء كتلة صورة لكل الشبكات و الشبكة التي تعطي قيمة عظمى لـــ SNR يتم اختيارها.نسمي هذه النظرة " SNR الأفضل " كما هو مبين في النتائج في الفقرة التالية.لقد تم اختيار معيار مقياس التعقيد بدقة.

أيضاً لتخفيض تأثير التشويش المسمى "رقعة الشطرنج " على الصورة المعاد تركيبها و تحسين نوعية وضوح الصورة، يُسمح بتداخل كتل الصورة المجاورة. و لقد أدركنا هذه الفكرة مع الأخذ بعين الاعتبار تداخل بنسبة 50% بين صفوف و أعمدة الكتل المجاورة.و من أجل كل بكسل في منطقة التداخل يتم حساب القيمة المتوسطة للبكسلات المتداخلة. و كما تظهر النتائج في الفقرة التالية فإن التداخل حسّن نوعية وضوح الصورة و زاد النسبة PSNR لها.

من بين الصور الأربعة في الشكل(3-4) لا يتم تقديم سوى صورة Lena في مجموعة التدريب و ذلك من أجل الطرق التكيّفيّة الثلاث. و أن شروط التوقّف لتعليم التوليد الخلفي في كل من التراكيب التكيّفيّة الثلاث (أي الإنتروبية و الفعالية و المعتمدة على النمط) هما العدد الأعلى للتكرار هو 10000 و عتبة الخطأ هي $\varepsilon = 1.0e-5$ من أجل العلاقة(3-3).



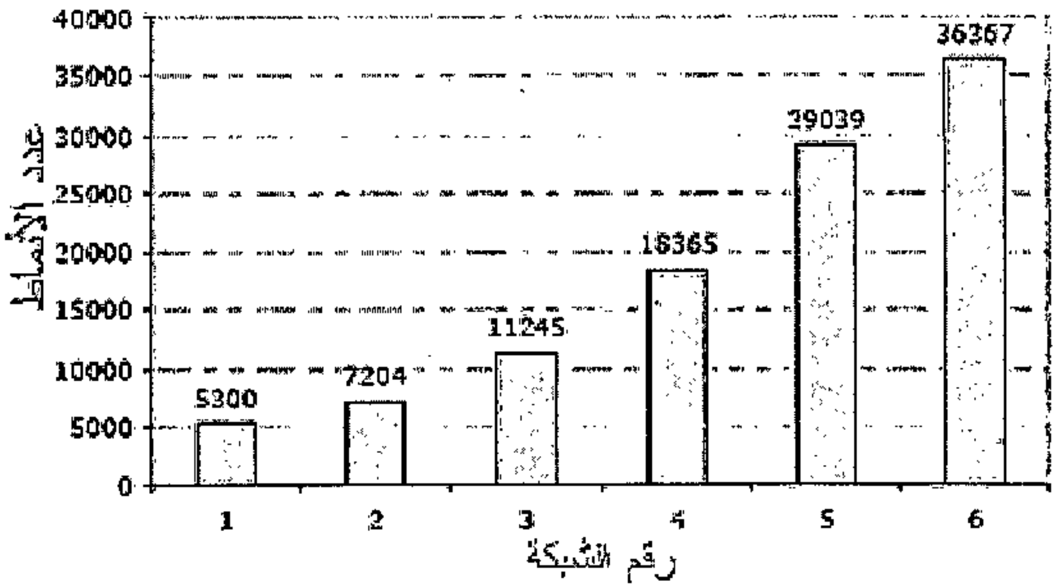
الشكل(3-6) بعض من عينات صور مجموعة التدريب المستخدمة في النظرة التكيّفيّة.

I) النتائج التجريبية باستخدام مقياس الإنتروبية:

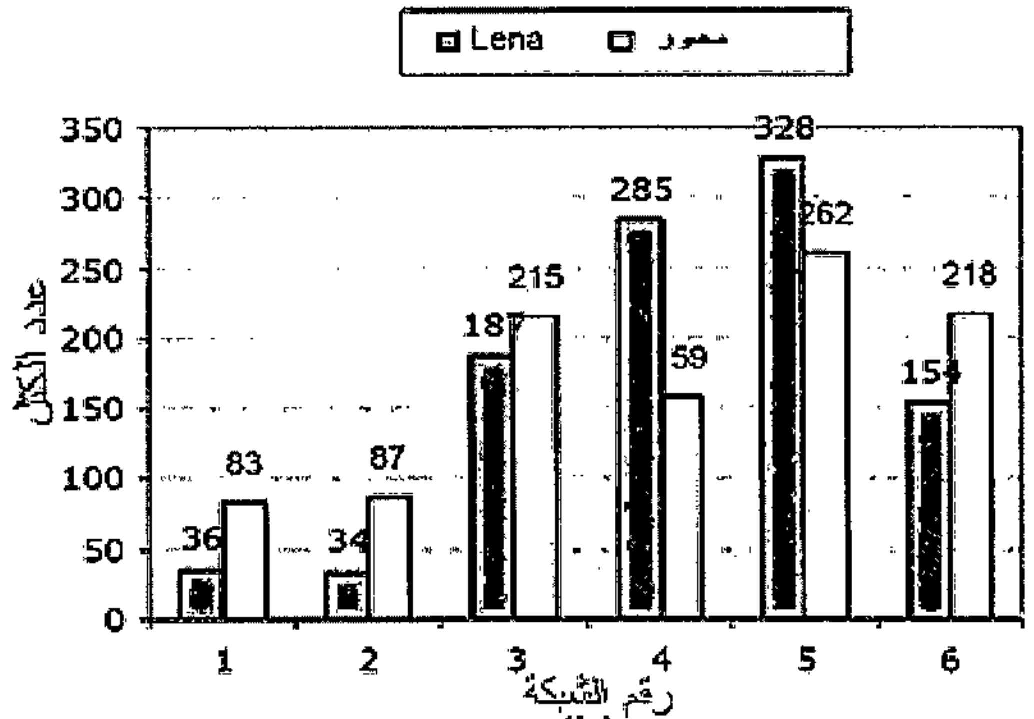
Experimental results using Entropy measure

يتم استخدام ست شبكات في معيار قياس تعقيد الإنتروبية و التي تعطي 3 بت إضافية لترميز عدد الشبكات (q في العلاقة(3-5)).و أن أعداد العصبونات المخفية لتلك الشبكات هي 4 و9 و 16 و 25 و 36 و 49 و التي تعطي نسب الضغط 16 و 7.1 و 4 و 2.5 و 1.7 على الترتيب.و بالتأكيد هناك احتمالات أخرى لاختيار عدد عصبونات الطبقة المخفية و لكن في كل منها يتطلب التعقيد الأعلى العدد الأعلى من العصبونات و العكس بالعكس.و أن الملاحظة الهامة في اختيار هذه الأعداد هي أن العدد الكبير من العصبونات المخفية ليس من الضروري أن يؤدي خطأ صغير من أجل مجموعة الاختبار على الرغم من أنه يعطي خطأ صغير في معطيات التدريب. و بالنظر إلى المدى المتوفر قيم مستوى الرمادية من 0 حتى 255 فإن قيمة الإنتروبية العظمى ستكون $log_2\ 256 = 8$. و يمكن الوصول إلى هذه القيمة العظمى عندما يكون لكل مستويات الرمادية القيمة نفسها. و بمعنى آخر تستقبل كتلة الصورة قيمتها العظمى من الإنتروبية عندما تحتوي العدد نفسه من كل قيم الرمادية في تلك الكتلة. و من أجل مستوى 256 في صورة فإن قيمة الإنتروبية العظمى ألا و هي 8 يمكن الوصول إليها عندما تحتوي كتلة الصورة على 256 بكسل على الأقل و لكن عند استعمالنا كتل صورة من الحجم 8 ×8 فإن هذه القيمة العظمى ستكون $log_2\ 64 = 6$.لقد استخدمنا العلاقة (3-13) لتخصيص الكتل للشبكات. و هي تعطي قيم رقم الشبكة في 6 ≥ NN ≤1.

$$\begin{cases} NN = Entropy + 1\ ; Entropy \neq 6 \\ NN = 6 \qquad\qquad ; Entropy = 6 \end{cases} \qquad (13\text{-}3)$$



[نمط 107520 = صورة (256 x256)105](a)



(b)

الشكل(3-7) :(a)المخطط الإحصائي لقيم الإنتروبية لجميع الصور في مجموعة التدريب و (b) لصورتي الاختبار Lena و المصور

إن اختيار عدد العصبونات المخفية يعتمد أيضاً على كون كتل الصورة ذات الإنتروبية الأعلى يجب أن ترمز بالشبكة ذات نسبة الضغط الأخفض.

يظهر الشكل(3-7-a) قيم NN من أجل جميع صور التدريب، وكما يظهر الشكل (3-7-b) قيم NN من أجل صورة Lena و المصور في مجموعة الاختبار. و يظهر توزيع عدد الأنماط لكل شبكة في معطيات التدريب كفاية المعطيات لتدريب الشبكات.

تم الوصول إلى شرط عتبة الخطأ في الطريقة التكيّفيّة المعتمدة على الإنتروبية قبل الوصول إلى العدد الأقصى للتكرار من أجل شرط التقارب في خوارزمية الارتداد الخلفي.و أن نتائج هذا التركيب التكيّفيّ مبيّنة في الجدول(3-2). لقد تمّ التقييم على صور الاختبار الأربع المبينة في الشكل(3-4) في كلا النظرتين " المعتمدة على التعقيد " و أفضل PSNR " مع و بدون تداخل من أجل كل صورة.يتم حساب نسبة الضغط (CR_a) باستخدام العلاقة(3-5).

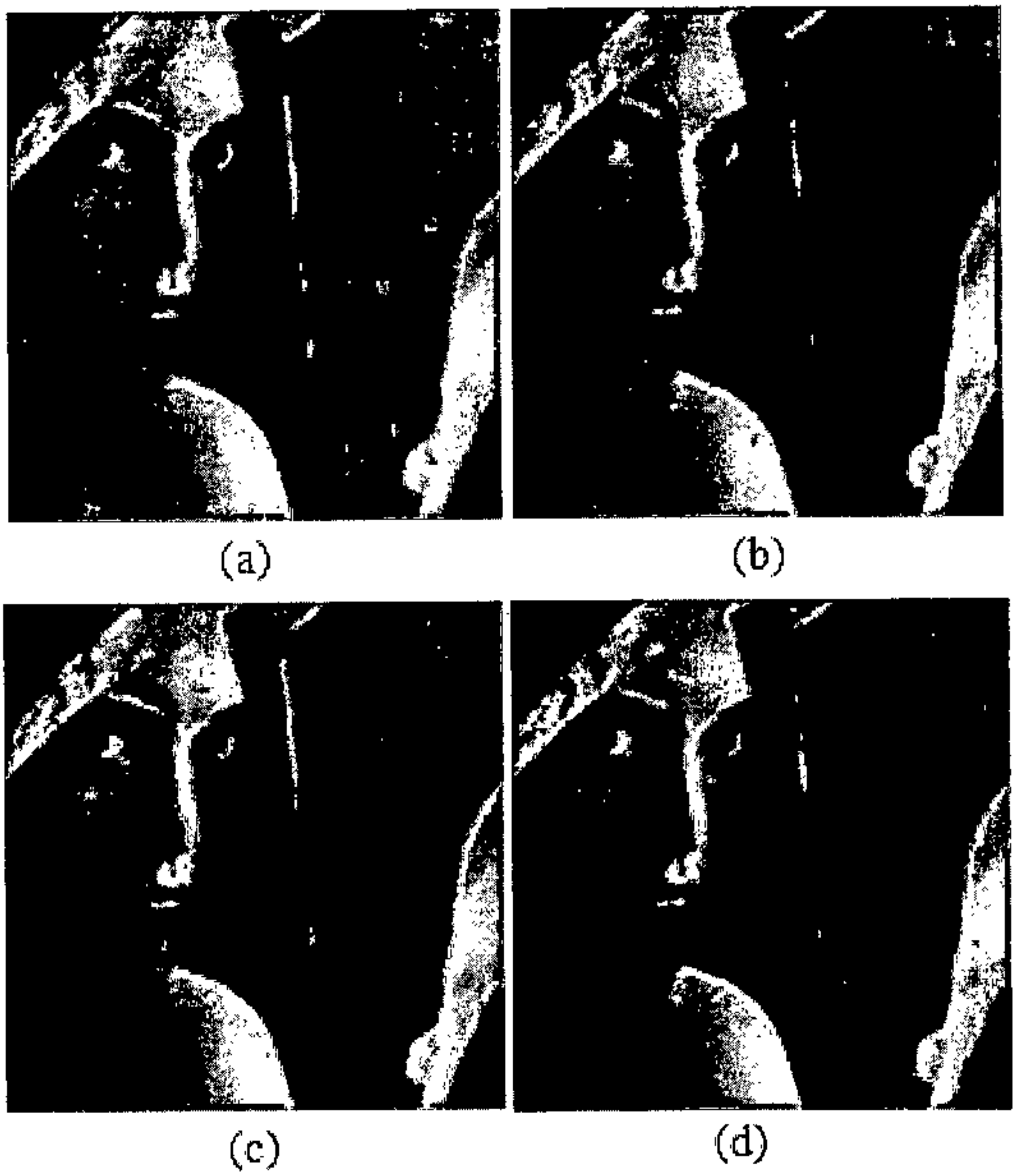
إن قدرة هذه الطريقة على إعادة تركيب الصور غير الموجودة في مجموعة التدريب كبيرة بشكل ملحوظ. و أن قيمة PSNR لكل صور مجموعة الاختبار متقاربة و بسبب قيمة تعقيدها المرتفعة لصورة الجمهور فإن قيمة CR لها أقل من الآخرين. و أن صور Lena المعاد تركيبها من أجل التجارب من (a) إلى (d) موضّحة بالشكل (3-8). و كما يظهر الشكل (a) (3-8) فإن تأثير رقعة الشطرنج واضح.إن استعمال تداخل الكتلة يخفض من تأثير هذا العيب كما هو موضح بالشكل (b) (3-8) و يزيد PSNR حوالي 1 dB دون التأثير على CR.إن اختيار شبكة " SNR الأفضل " من أجل كتل الاختبار يُحسّن بشكل ملحوظ قيمة PSNR و النوعية البصرية كما تُظهر الأجزاء (c) و (d) في الجدول (3-2) و الشكلين (c) (3-8) و(d)(3-8).

الجدول(3-2) نتائج الضغط بواسطة تركيب BPNN التكيّفيّ باستخدام معيار مقياس تعقيد الإنتروبية

الشبكة المختارة	صورة الاختبار	PSNR (dB)	CR_a
(a) الغير متداخلة المعتمدة على التعقيد	Lena	35.50	2.76
	مصور	35.11	2.83
	جمهور	37.92	1.71
	فلفل	35.62	2.20
(b) المتداخلة المعتمدة على التعقيد	Lena	36.70	2.76
	مصور	36.57	2.82
	جمهور	38.68	1.71
	فلفل	36.25	2.20
(c) الغير متداخلة ذات أفضل SNR	Lena	43.10	1.79
	مصور	39.24	2.04
	جمهور	38.11	1.70
	فلفل	38.06	1.78
(d) المتداخلة ذات أفضل SNR	Lena	44.55	1.79
	مصور	41.04	2.06
	جمهور	39.73	1.69
	فلفل	39.71	1.74

الجدول(3-3) معدل التعقيد(الإنتروبية) من أجل صور مجموعة الاختبار (صور 256× 256 :1024 كتلة من الحجم 8× 8

الصورة	جمهور	فلفل	Lena	مصور
معدل الإنتروبية	5.32	4.52	3.71	3.47



الشكل(3-8) صورة Lena المعاد تركيبها باستخدام نظرة الضغط التكيفية المعتمدة على الإنتروبية باستخدام أربع طرق ضغط مختلفة من (a) حتى (d) من الجدول(3-2)

و من وجهة نظر أخرى فإن عيوب هذه النظرة أنها تزيد زمن CODEC و تنقص CR.يُظهر انخفاض CR بأن الشبكات ذات CR منخفضة يتم اختيارها لترميز الكتل بتعقيد منخفض.و أن الملاحظة البارزة للنتائج من (a) حتى (d) في الجدول(3-2) تبرز تحسّن PSNR بين الصورتين Lena و الجمهور. و السبب في هذا هو الاختلاف في مستوى التعقيد

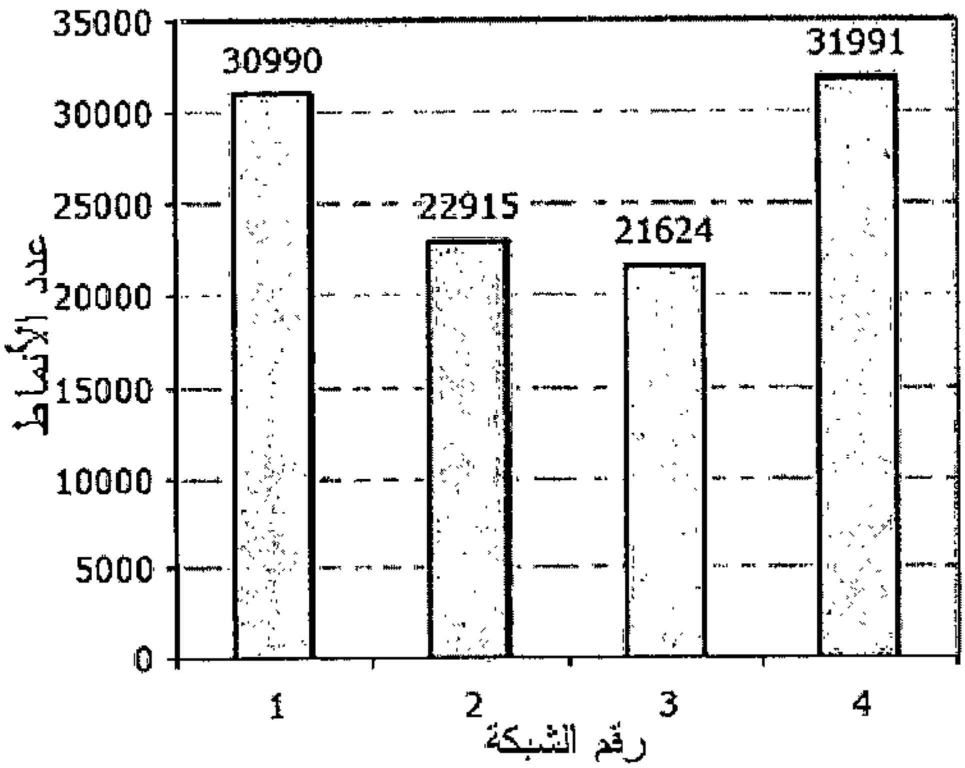
بين هذه الصور. و أن قيم معدل التعقيد(الإنتروبية) في مجموعة الاختبار التي قمنا باستخدامها موضحة بالجدول(3-3).

إن كلاً من الصورتين Lena و المصوّر أقل تعقيداً من صورتي الجمهور و الفلفل. و أن صورة الجمهور هي الأكثر تعقيداً و لكن استخدام " SNR الأفضل " لم يُزوّد بتحسين هام بنسبة PSNR لصورة الجمهور المُعاد تركيبها. إن هذا بسبب التحسين الناتج عن طريقة " أفضل SNR " و خصوصية هذه الطريقة في استخدام الشبكات العصبونية ذات CR المنخفضة من أجل كتل ذات تعقيد منخفض في صورة ما. و في صورة الجمهور فإن 85% من الكتل لها تعقيد بحيث NN = 6 و 5 % فقط بحيث يكون NN ≤ 4 . و هذا يعني أنه يمكننا أن نستخدم شبكات ذات CR منخفضة فقط من أجل 5 % من الكتل. و لذا يمكن تبرير أي تخفيض على CR لهذه الصورة بالمقارنة مع الصور الأخرى.

II) النتائج التجريبية باستخدام مقياس الفعالية

Experimental results using Activity measure

في هذه الطريقة تم استخدام 4 شبكات حجم كتلة الدخل لها هو نفسه كما في طريقة الإنتروبية. و يكون الشبكات 9 و 16 و 25 و 36 عصبون مخفي و التي تعطي نسب ضغط CR_s هي 7.1 و 4 و 2.5 و 1.7 على الترتيب.نظرياً فإن القيمة العظمى للفعالية في العلاقة(3-8) من أجل كتلة صورة ذات المستوى 256 و الحجم 8 ×8 هي أعلى من 3.0e + 6 و لكن هذه القيمة من أجل مجموعة التدريب التي قمنا باستخدامها هي حوالي 1.9e + 6 أي هناك عدد صغير جداً من الكتل و التي لها فعالية مرتفعة قريبة من هذه القيمة. و لقد اخترنا 4 مجالات مختلفة هي]0,465] و]465.4073] و]4073,20154] و] ∞,20154] من أجل كل شبكة.يتم هذا بالنظر إلى جميع قيم الفعالية في مجموعة التدريب.يظهر الشكل(3-9) عدد الأنماط في مجموعة التدريب للشبكات الأربع تلك.



الشكل(3-9) المخطط الإحصائي لعدد كتل التدريب من أجل مجالات قيم الفعالية الأربع من أجل 105 صورة في مجموعة التدريب(بمعنى عدد الأنماط = (حجم الصور :265 × 256) 105 = 107520)

يظهر الجدول(3-4) نتائج الطريقة التكيفيّة. إن PSNR الناتجة عن هذه الطريقة هي أفضل من الإنتروبية من أجل الصور الأقل تعقيداً في الاختبارين (a) و (b). و تظهر هذه النتائج أن طريقة المقياس هذه تقوم بحساب التعقيد بشكل أفضل من الإنتروبية فقط من أجل الصور الأقل تعقيداً.

الجدول(3-4) نتائج الضغط بواسطة تركيب BPNN التكيّفيّ باستخدام معيار مقياس تعقيد الفعالية

الشبكة المختارة	صورة الاختبار	PSNR (dB)	CR_a
(a) الغير متداخلة المعتمدة على التعقيد	Lena	42.04	11.94
	مصور	37.79	9.25
	جمهور	36.15	4.78
	فلفل	34.28	6.93
(b) المتداخلة المعتمدة على التعقيد	Lena	43.12	11.79
	مصور	38.24	9.16
	جمهور	36.41	4.75
	فلفل	34.93	6.67
(c) الغير متداخلة ذات SNR الأفضل	Lena	43.99	7.96
	مصور	39.56	7.47
	جمهور	38.58	4.46
	فلفل	35.43	4.82
(d) المتداخلة ذات SNR الأفضل	Lena	45.09	7.93
	مصور	39.04	7.38
	جمهور	36.82	4.38
	فلفل	35.74	4.81

أن العدد الأصغر للشبكات في هذه الطريقة هو السبب في أن " SNR الأفضل " لا تزود بالنسبة PSNR بالجودة ذاتها كما في الطريقة المعتمدة على الإنتروبية. و من ناحية أخرى فإن هذه الأعداد من الشبكات تُعطي 2 بت إضافي فقط بالمقارنة مع 3 بت في الطريقة المعتمدة على الإنتروبية. مع الأخذ عين الاعتبار اختيار عدد العصبونات في الطبقة المخفية، فإن هذه الطريقة تُزوّد بأفضل CR. و تُشير نتائج هذه الطريقة في الحالة (d) بأن طريقة " أفضل SNR " للصور المُعقدة (الجمهور و الفلفل) لا تُظهر مستوى عال من التعقيد. و سبب ذلك هو تعقيد هذه الصور.

III) النتائج التجريبية باستخدام المقياس المعتمد على النمط

Experimental results using Pattern-based measure

تم محاكاة هذه الطريقة باستخدام 16 شبكة، شبكة واحدة لكل نمط في الشكل(3-3). و تتعلق CRs بهذه الأنماط بحيث تكون 16:1 للأنماط 1 و 2 ، و 4:1 للأنماط من 6 إلى 8 و من 9 إلى 12 و 7.1:1 للأنماط من 13 حتى 16 و 2.5:1 للأنماط 3 و 4.إن اختيار عدد العصبونات في الطبقة المخفية يعتمد على الطرق البصرية و الحدسية. و يُمكن أن نفرض الأنماط المضغوطة في الطبقة المخفية كنسخة مُصغرة للأنماط الأصلية و نستخدم النسبة CR لكل الشبكات هي 64:16. و أن نتائج CODEC لهذه الطريقة موضحة بالجدول(3-5).

إن هذه النتائج غالباً في نفس الاتجاه من أجل الطريقتين السابقتين.التداخل يزيد PSNR حوالي 1 dB و طريقة اختيار شبكة " SNR الأفضل " تُعطي نوعية عالية للصور المُعاد تركيبها حتى للصور المُعقدة.و أن هذا بسبب العدد الأعلى للشبكات المُستخدمة في هذه الطريقة بالمقارنة مع الطرق السابقة. أيضاً و بسبب تركيب الشبكة فإن CR للصور هي أعلى من التركيب المُعتمد على الإنتروبية و أخفض من المُعتمد على الفعالية. و يتوفر هناك 4 بت إضافية في هذه الطريقة لترميز رقم الشبكة.بالإضافة إلى ذلك فإن تخفيض قيمة CR باستخدام " SNR الأفضل " معقول بسبب العدد الأعلى من الشبكات.

الجدول(3-5) نتائج الضغط بواسطة تركيب BPNN التكيّفيّ باستخدام معيار النمط المعتمد على التعقيد

CR_a	PSNR (dB)	صورة الاختبار	الشبكة المختارة
6.67	39.11	Lena	(a) الغير متداخلة المعتمدة على التعقيد
6.69	34.44	مصور	
5.11	30.66	جمهور	
6.11	32.01	فلفل	
6.67	40.20	Lena	(b) المتداخلة المعتمدة على التعقيد
6.73	34.80	مصور	
5.10	32.00	جمهور	
6.19	32.86	فلفل	
3.24	44.00	Lena	(c) الغير متداخلة ذات SNR الأفضل
3.70	38.60	مصور	
2.90	36.91	جمهور	
3.00	36.08	فلفل	
3.26	46.76	Lena	(d) المتداخلة ذات SNR الأفضل
3.71	40.45	مصور	
2.89	39.41	جمهور	
3.01	38.14	فلفل	

C) مُقارنة بين الطرق المُعتمدة على الإنتروبية و الفعالية و النمط

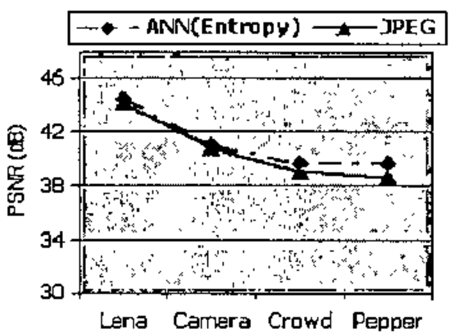
Comparison of Entropy, Activity and Pattern-based Methods

قارنّا فيما يلي بين ثلاث طرق تكيّفيّة معاً و من ثم قارنّا هذه الطرق في أحسن حالاتها، و من ثم " SNR الأفضل" مع المتداخلة و مع JPEG القياسيّة. كما تُظهر النتائج في الجداول(3-3) و (3-4) و (3-5) ، تُعطي الإنتروبية تقدير جيد للتعقيد من أجل الصور المعقدة. و من جهة أخرى فإن الفعالية تُنجز أفضل تقدير لأبسط الصور و الأقل تعقيداً. فقد

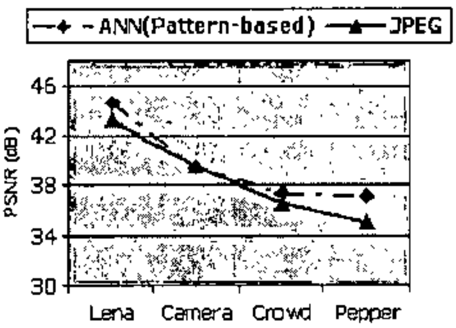
أعطت هذه الطريقة CRs الأعلى بسبب تركيب شبكتها. و أن طريقة معيار التعقيد المُعتمد على النمط أعطى PSNR ذاتها و CR أعلى من الإنتروبية و CR أخفض من الفعالية.

من أجل أفضل مُقارنة بين هذه الطرق و الأخذ بعين الاعتبار كل من PSNR و CR معاً، فإننا نقوم بمقارنتها مع JPEG القياسية. و النتائج مبينة بالشكل(3-10).إنه يُوضح قابلية طرق الضغط التكيّفيّة بالمُقارنة مع خوارزمية JPEG القياسية. و تتم هذه المقارنة بنسبة البت ذاتها من أجل كل طريقة. تُظهر هذه النتائج تفوق طرق الضغط المعتمدة على ANN.

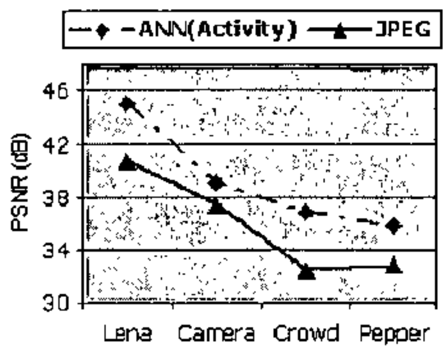
تختلف الطرق التكيّفيّة الثلاث إضافة لنتائج الضغط من وجهة نظر المبدأ الأساسي.و أن الإنتروبية هي مقياس لارتياب متغير عشوائي و الذي يُحدد كمية المعلومات لمصدر ما مثل الصورة.



(a)



(b)



(c)

الشكل(3-10) مقارنة الطرق التكيّفيّة مع طريقة ضغط JPEG القياسية(a) الإنتروبية ،(b) الفعالية ،(c) المعتمدة على النمط

إن الفعالية هي طريقة حدسية لتقدير تعقيد كتلة صورة باستخدام الاختلاف بين كل بكسل و البكسلات التي تجاوره. ولا تُميز هاتان الطريقتان مكان أو اتجاه التعقيد و تُعطي قيمة متوسطة للتعقيد. فعلى سبيل المثال قيم التعقيد المقيسة باستخدام الإنتروبية أو الفعالية للأنماط 15 و 16 في الشكل(3-3) هي ذاتها، إن هذين النمطين مختلفين تماماً من وجهة نظرة تعليم ANN.و من جهة أخرى، فإن الطريقة المعتمدة على النمط تُصنف كتلة الصورة ضمن بعض الأنماط المُعرفة مسبقاً. و في الحقيقة هذه الطريقة ليست معيار قياس تعقيد و تُشير إلى مشكلة أخرى تختلف عن المشكلة التي تُشير إليها الطرق الأخرى. و أنها لا تأخذ بعين الاعتبار قيم تدرج الرمادي بالضبط و تُمثل كل من الكتل 4 × 4 الجزئية في كتلة واحدة أو نمط أبيض.و أن الحل الوحيد لهذه المشكلة هو دمج الطريقة المعتمدة على النمط مع الطرق الأخرى.

5-3 الخلاصة Conclusion:

لقد استعرضنا استعمال شبكة Preceptoron العصبونية المتعددة الطبقة لضغط الصور. و لكونها لم يتم الحصول على نتيجة مقبولة للضغط باستخدام شبكة واحدة، تم استخدام نظرة تكيّفيّة.إنها تستخدم شبكات مختلفة من أجل كتل صورة مختلفة بالنظر لقيم تعقيدها.تم تقديم و تقييم ثلاث طرق هي مقياس الإنتروبية و الفعالية و المعتمد على النمط. و لقد أظهرت نتائجنا التجريبية الحصول على أفضل نوعية بصرية من خلال تداخل كتل الصورة المتجاورة. أيضاً فإن اختيار صور وفق معيار " الأفضل SNR " بدلاً من استخدام معيار التعقيد يُزوّد بأعلى نوعية صورة و أفضل PSNR. و العدد الأعلى من الشبكات يُزوّد بأفضل أداء في نظرة " أفضل SNR " و لكن ذلك سيؤدي إلى أخفض CR. و من ناحية أخرى، التداخل و اختيار الشبكة يحتاج إلى المزيد من التحقّق للحصول على أفضل نوعية للصورة المُعاد تركيبها. و تُظهر مقارنة النتائج مع خوارزمية JPEG القياسية أداء أفضل لطريقتنا في كل من مقياس PSNR و النوعية البصرية.

في هذا الفصل أعداد الشبكات التي تُزوّد بنسب ضغط مختلفة لم تُحسّن. فإن هذا مُتوقع لأن اختيار العدد الأكبر من الشبكات و اختيار نسب الضغط المثلى للشبكات يُزوّد بنتائج أفضل. و لهذا الغرض بعض أنواع الشبكات العصبونية مثل "الاتصال المتسلسل " يُمكن أن تستخدم. في هذه الشبكات يُمكن أن نختار عدد العصبونات في الطبقة المخفية بالطريقة التي يتم

من خلالها الحصول على نسبة الضغط المثلى. و من بين الطرق الثلاث المقترحة فإن الإنتروبية و الفعالية لا تستخدم تكييف الأنماط و المعتمدة على النمط مستويات الرمادي بشكل ملائم. و أن أخذ قيم مستويات الرمادي في الطريقة المعتمدة على النمط بعين الاعتبار يمكن أن يُعطي نتائج أفضل. و يتم تدارك ذلك بدمج الفعالية و الإنتروبية مع المعتمدة على النمط. و مع ذلك يبدو أن الطرق المعتمدة على ANN ليست مرنة كفاية في التحكم بنسبة الضغط، و لكن من الممكن أن يكون لديها مجموعة من الشبكات المُدرّبة بنسب ضغط مختلفة بدلاً من شبكة واحدة في كل حالة.

3-6 النظرة المستقبلية:

يمكن في مستقبلاً كتطوير لبحث ضغط الصور باستخدام الشبكات العصبونية أن يُأخذ بعين الاعتبار بناءها بحيث يكون عدد الطبقات المخفية أكبر و الذي يسمح للشبكة الحصول على المزيد من التدريب التفاعلي و التعلّم الأكثر تطوراً. و وفقاً لذلك يُمكن تطبيق و تطوير خوارزميات للضغط ذات أداء عالي على تلك الأنواع من الشبكات العصبونية.

الفصل الرابع

تصفية الصور باستخدام الشبكات العصبونية

1-4 مقدمة Introduction:

تحدد أنواع عديدة من التشويهات جودة الصور الرقمية أثناء استلامها و تشكيلها و تخزينها و إرسالها. و غالباً تفسد الصور بسبب التشويش.و أن شدة التشويش إما أن تكون مرتفعة أو منخفضة نسبياً و في الحالتين فإنها تؤدي إلى فقدان تفاصيل الصورة . و من المهم إزالة التشويش من الصور قبل استعمالها في تقنيات معالجة الصور مثل كشف الحواف و التقسيم و الضغط ..الخ . فقد تم في الماضي اقتراح العديد من طرق التصفية لمعالجة عملية إزالة التشويش . و أن مصفي المتوسط القياسي هو واحد من أشهر مصفيات الصور ذات التدرج الرمادي و الذي يصنف درجات كثافة البكسلات و يستبدل البكسل المركزي بقيمة المتوسط . و أن توسيع فكرة مصفي المتوسط لتشمل الصور الملونة ليست بالأمر السهل لأنه قد يحدث تشوه اللون عند تطبيق هذا المصفي بشكل منفصل على كل مكون من مكونات متجهات اللون. و عند تطبيق مصفي المتوسط غير الموجه(Scalar Median Filter)بشكل منفصل على كل مكون مفرد من متجهات اللون يمكن أن يحدث تشوه اللون. فتم اقتراح الطريقة المسماة مصفي متوسط المتجه (Vector Median Filter) أو VMF، و الذي يأخذ بعين الاعتبار مكونات اللون الثلاث[41] .عدة تعديلات قد تم تطبيقها على مصفي متوسط المتجه مثل مصفي المتوسط الإتجاهي (Directional Median Filter)[67]، و مصفي متوسط المتجه المركزي المرجح(Central Weighted Vector Median Filter)[69]. كما تم تقديم فكرة تطبيق مصفي على فضاء لون مختلف عن RGB [68] و [42] حيث تم استخدام فضائي اللون HSI و L*a*b على الترتيب.

إن عيب نظرات متوسط المتجه التقليدية تكمن في تطبيق المتوسط على كل بكسل بغض النظر عن تعرضه للعطب أم لا.إن الحل الذهني للتغلب على هذا العائق هو أولاً اكتشاف البكسلات المعطوبة و من ثم تطبيق التصفية على هذه البكسلات فقط. و اقترحت عدة خوارزميات تبعاً لطريقة الخطوتين تلك[72][70][64][53] .و أن أحد المشاكل الرئيسية في كشف التشويش هي الصعوبة في التفريق بين الحافة و التشويش. و حيث أنه في فضاء الكثافة كلاً منهما تبدو كذروة في جوارها. ففي [53] إذا كانت قيمة الرمادية للبكسل المركزي أكبر من عتبة معينة فإنه يعتبر كتشويش. و أن عيب هذه الطريقة يكمن في كون نسبة الخطأ

الإيجابي مرتفعة جداً و أغلب الحواف يتم اكتشافها كتشويش. و علاوة على ذلك، ففي [64] تم التعامل فقط مع قيم التشويش العشوائية و في [72] مُكلف حسابياً. و في [73] تم تدريب الشبكة العصبونية على كشف التشويش.و لكنها تعمل فقط على قيم التشويش الثابتة و يتم تعليم الشبكة بطريقة الإشراف.و في [59] تم استخدام التقنية الضبابية فإن دالة كانت ذات تأثير كبير على أداء خطوة كشف التشويش. و تعتمد هذه الدالة على عدة متغيرات التي يجب تعديلها.

هدفنا هو تطوير تقنية سريعة و تحتاج إلى تعديل أقل عدد ممكن من المتغيرات. و هنا نقترح في هذا الفصل حل ذو خطوتين لإزالة التشويش.هاتان الخطوتان هما:

1- كشف التشويش بالاعتماد على شبكة هوبفيلد العصبونية.

2- تصفية التشويش بالاعتماد على VMF في اثنين من فضاءات اللون.

إن كشف التشويش يتم بدون بإشراف لكون HNN شبكة تعلّم ذاتي و تحتاج لتعديل أقل عدد من المتغيرات.بتعديل بسيط سوف تعمل المتغيرات جيداً على مجموعة واسعة من المتغيرات. و تم تطوير طريقة تصفية التشويش بالتغلب على العيب في تطبيق VMF على مجال وحيد اللون. و اختبرت الخوارزمية المقترحة على صور مختلفة عبر مستويات مختلفة من التشويش و قد أظهرت أداء ممتاز كماً و نوعاً في كلا المقياسان PSNR و " متوسط الخطأ المطلق " (Mean Absolute Error).

يصف المقطع الثاني خوارزميتنا لكشف التشويش.يصف المقطع الثالث خوارزمية مصفي التشويش المُقترح. و يُقدم المقطع الرابع النتائج التجريبية متبوعاً بمناقشة و ملاحظات ختامية في المقطعين الخامس و السادس.

4−2 كشف التشويش Noise Detection:

إن شبكة هوبفيلد العصبونية (HNN) ذات تركيب وحيد الطبقة بتغذية عكسية أو هي شبكة عصبونية تكرارية حيث ترتبط كل وحدة بالأخرى. و هي تأخذ القيم الثنائية كدخل و بمعنى آخر إما −1 و 1 أو 1 و 0 و تعطي قيم ثنائية كخرج بالاعتماد على دالة التنشيط. و كل ارتباط بالشبكة يمكن أن يتعلق بوزن و يُمثّل بواسطة مصفوفة ثنائية تدعى مصفوفة

الوزن W.يتم تدريب شبكة هوبفيلد بتغذيتها بدخل مكون من متجهات السمة أو متجهات النمط المتوافقة مع أصناف مختلفة. و من ثم تُستعمل الشبكة لتصنيف أنماط الاختبار ضمن الأصناف و التي كانت قد تدرّبت بها.و يُظهر تركيب هذه الشبكة بالشكل(4-1). و لقد تم تعديل HNN[44]، و تم استخدامها لكشف الحواف في الصور ذات التدرج الرمادي. و لقد كيّفنا شبكة HNN المُعدّلة لكشف التشويش. ليكن x شعاع دخل W، مصفوفة الوزن، a دالة التنشيط $(a= Wx_t)$ و لتكن $sgn[y]$ تُمثّل دالة سيغمويند .

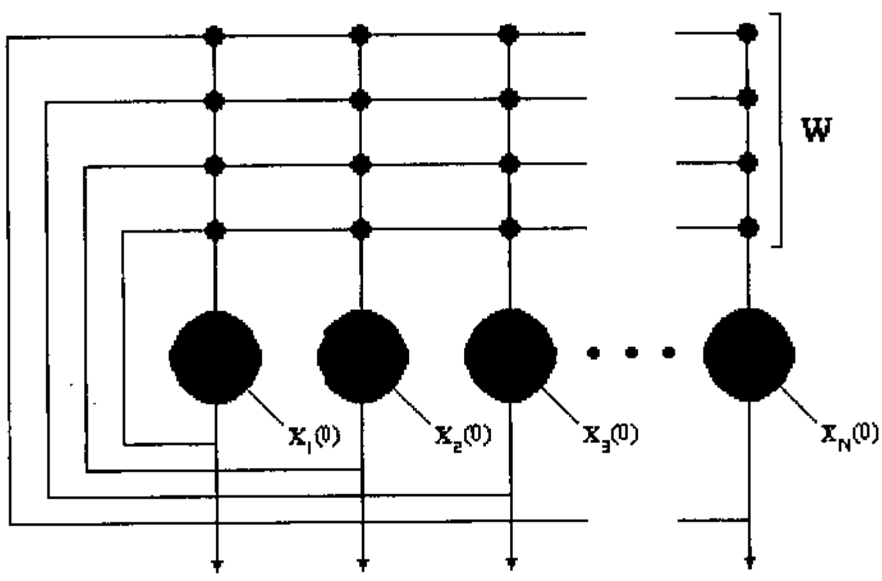
و تكون قاعدة الإرسال هي:

$$x_{t+1} = sgn[Wx_t] \text{ , t = 1...T} \qquad (1\text{-}4)$$

$$sgn[a_j] = \begin{cases} a_j = 1\,, & If\ a_j > 0 \\ a_j = -1\,, & If\ a_j < 0 \\ Else & a_j = 0 \end{cases} \qquad (2\text{-}4)$$

و قاعدة التعلّم هي:

$$W_{t+1} = W_k + \eta[x_0 x_o^T - (Wx_t)(Wx_t)^T] \qquad (3\text{-}4)$$



الشكل(4-1) تركيب HNN

حيث η ($0 < \eta$) تُمثّل مُعامل التعلّم و t ($0 < t$) عدد مرات تكرار متجه الحالة و الذي يجب إنجازه قبل تحديث مصفوفة الوزن.

4-2-1 متجهات السمة Feature Vectors:

تُشتق متجهات السمة من أجزاء الصورة المحيطة بالبكسلات المختارة عشوائياً. و يتم الأخذ بعين الاعتبار البكسلات الثمان المجاورة لكل بكسل مختار أو N8 و يعطي متجه سمة مكون من تسع أبعاد. و على الرغم من ذلك فإن HNN القياسية تأخذ فقط القيم الثنائية، فلقد قمنا بتعديلها لتأخذ القيم التي تقع في المجال [1,1-] و ذلك بإعادة قياس قيم متجه السمة. و تقوم دالة السيغموند بعد تكرار واحد بجعل كل القيم ثنائية. و يُمكن لشبكة مُدرّبة بواسطة متجهات السمة هذه أن تكشف البكسلات المشوشة تلك و التي تختلف تماماً عن N8. و لكي تُميّز HNN حتى البكسلات المشوشة تلك و التي تختلف بالحد الأدنى عن N8 يتم إعادة قياس متجهات الدخل بشكل تكراري لتُركز الشبكة على مدى صغير من القيم في كل مرّة. و يتم إعادة قياس متجهات الدخل بالعلاقة التالية:

$$x_r = x_0 - 1 + s \times t \qquad (4\text{-}4)$$

حيث x_r تُمثل الدخل المعاد قياسه، s($0 < s \leq 1$) مجال خطوة إعادة القياس و t هو مجال اختلاف الزمن المتقطع(t من 0 حتى $2/s$). و على سبيل المثال إذا أردنا مدى $s = 0.2$ بين كل خطوة زمن فإن المتغير t سيكون 10 , ... ,1 ,0. و هكذا فإن شعاع الدخل نفسه سوف يظهر 11 مرة ، و في كل مرة من نقطة قياس مختلفة.

4-2-2 تدريب الشبكات Training the Networks:

باعتبار أن متجهات السمة هي أجزاء عشوائية صغيرة من الصورة، فإن صورة وحيدة ستؤدي غرض التدريب.و لأننا نريد من HNN أن تكشف الاختلاف بين التشويش و الحواف و المناطق الناعمة، فإن أي صورة تحتوي على حواف متنوعة و تشويش كافي تكون مناسبة للتدريب. و هكذا يتم تدريب HNN بمتجهات الدخل التي تقابل ثلاث أصناف هي التشويش و الحواف و المناطق الناعمة.

4-2-3 تعريف نمط التشويش Identification of the noise pattern:

أنه من الضروري قبل اختبار الشبكة فهم سلوكها في حال وجود أو عدم وجود التشويش.بسبب وجود 9 عصبونات في HNN التي اخترناها فإننا نحصل على 9 عصبونات في الخرج. و سوف تستجيب الشبكة ليس فقط للتشويش في الصورة بل للمناطق الناعمة و للحواف أيضاً.

إذا أُعتبر البكسل المركزي ضمن جوار ناعم بمعنى آخر إذا كانت جميع البكسلات في نافذة التصفية غير مُشوشة فإن العصبونات التسع جميعها سوف تسلك نفس السلوك و سيكون للخرج نفس القيمة أي إما +1 أو −1.

أما إذا كان البكسل المركزي معطوب فإن سوف يسلك بشكل واضح سلوكاً مختلفاً عن N8 ، و بمعنى آخر إذا كانت N8 جميعها تُعطي +1 كخرج، فإن خرج البكسل المركزي سيكون −1 و العكس بالعكس كما هو موضح:

1+	1+	1+	1+	1−	1+	1+	1+	1+

أو

1−	1−	1−	1−	1+	1−	1−	1−	1−

من المرغوب الحصول على خرج وحيد ، فعلى سبيل المثال +1 للبكسل المشوش و 0 لغير المشوش. و للحصول على ذلك ترتبط العصبونات التسع جميعها بعصبون وحيد هو عصبون الخرج.و تكون دالة عصبون الخرج كما يلي:

$$Output = \begin{cases} 1 & if\ \sum_{(i=1, i\neq 5)}^{9} x_i = 8\ and\ x_5 = -1, or \\ & \sum_{(i=1, i\neq 5)}^{9} x_i = -8\ and\ x_5 = 1 \\ 0 & otherwise \end{cases} \qquad (5\text{-}4)$$

و هكذا فإن إعطاء صورة اختبار لشبكة HNN مُدرّبة يُصنّف ثنائياً كل بكسل إلى أحد صنفين إما تشويش أو عدم تشويش.

4-2-4 توسيع طريقة متجهات السمة إلى الصور الملوّنة

Extending the Feature Vectors method to color images

إن طريقة متجهات السمة التي تم مناقشتها هي للصور ذات التدرج الرمادي.من أجل توسيعها لتشمل الصور الملونة نقوم بتطبيقها على كل قناة بشكل مستقل. و إذا تم تصنيف بكسل كتشويش في أحد القنوات الثلاث R أو G أو B فإنه يُصنّف كبكسل معطوب في الصورة الملونة.

3-4 تصفية التشويش Noise Filtering:

بشكل عام، يتم حساب متوسط المتجه بأخذ مجموع المسافات بين أي بكسل و آخر في جواره ضمن نافذة التصفية، و البكسل الذي يُعطي الحد الأدنى من تلك المسافات يؤخذ ليكون متوسط المتّجه لنافذة التصفية تلك.

إن العيب الرئيسي لمُصفي متوسط المتجه هو ذلك حيث لا يوجد حد أدنى واضح. و هذا يجعل أكثر من بكسل ليكون ملائم لاختياره كمتوسط متجه. و لتكن N المجموع الكلي و S تُمثل المجموعة الجزئية البكسلات و الملائمة لتكون متوسط المتجه في نافذة التصفية المدروسة.فإذا تم اعتبار المنطقة منطقة صقيلة، فإن اختيار أي بكسل من المجموعة S سوف لن يُحرّف الإدراك البصري للصورة لأن كل منها سيكون لديه قيمة أكبر أو أقل.و لكن في الحالة التي تكون فيها المنطقة غير صقيلة، فإن اختيار أي بكسل من المجموعة S بدون تحقّق يُمكن أن يُعطي نتائج غير مقبولة لأنه يمكن أن لا يكون للبكسلات القيّم ذاتها. و لذا يجب أن يكون هناك معيار لحل هذا الالتباس.إن VMF القياسي لا يفرض أي قاعدة لانتقاء أفضل بكسل ملائم في هذه الحالة[41].

لذا نقترح VMF المُحسّنة و التي تعتمد على الحقيقة بأن تشويهات اللون في الصورة لها تأثير مباشر على إدراك اللون بالنسبة للمشاهد. و وفقاً لذلك يجب أن يكون هناك معايير إدراك إضافية لاختيار أفضل مُرشح في نافذة التصفية.

تُعتبر درجة اللون كمقياس وحيد لاختبار اللون وتُمثّل الاتجاه الأكثر حساسية في الفضاء ثلاثي الأبعاد في تحليل الصورة الملونة.لذا، في فضاء HSI [68] تُنجز VMF في فضاء اللون HSI فقط وفق القواعد التالية:

1- التغيّر في قيم درجة اللون يجب أن أقل ما يُمكن.

2- التغيّر في الإشباع يجب أن يكون أقل ما يُمكن.

نحل الالتباس في خوارزميتنا المقترحة بأخذ VMF أولاً في الفضاء RGB و من ثم الانتقال إلى الفضاء HSI. و تُبرز تفاصيل الخوارزمية كما يلي:

خوارزمية VMF المُحسنة:

الخطوة 1: حساب متوسط المتجه في الفضاء RGB في نافذة التصفية.

الخطوة 2: جمع كل البكسلات الملائمة لتكون متوسط نافذة التصفية.

الخطوة 3: إذا كان هناك أكثر من بكسل واحد ملائم لكي يكون متوسط، عندها يتم اختيار البكسل الذي يتقارب من وسط درجة اللون في الفضاء HSI.

الخطوة 4: إذا كان هناك أكثر من بكسل واحد ملائم للشرط في الخطوة 3 عندها يتم اختيار البكسل الذي يتقارب من متوسط الإشباع في الفضاء HSI.

4-4التجارب و النتائج Experiments and Results:

تم تدريب HNN باستخدام مُعامل التعلّم$0.01 = \eta$ مع 2000 اختيار عشوائي لمتجهات السمة. و يتم اختيار بكسل عشوائي في صورة التدريب و كما يتم استخدام نافذة جزئية من الحجم 3×3 من حوله كمتجه سمة لتدريب الشبكة.و بكون اختيار البكسل يتم عشوائياً، فإن ذلك يضمن بأن الشبكة ستُدرّب بعدد كاف من البكسلات المُشوّهة.أي صورة ذات تدرج رمادي تعرضت لتشوه أكبر من 5% من التشويش ستضمن ذلك التدريب أيضاً.

و قد أُستخدمت الصورة الموضحة بالشكل(4-2) لتدريب HNN. فهي صورة ذات تدرج رمادي و تم تشويهها بـــ 10 % بالتشويش المسمى " ملح و فلفل" . و لدى هذه الصورة تشكيلة جيدة من الحواف بالإضافة إلى التشويش و بالتالي فإن HNN تتعلم الفرق بين الحواف و التشويش.



الشكل(4-2) صورة ذات تدرج رمادي مشوهة بتشويش " ملح و فلفل" و المُستخدمة في تدريب HNN

نماذج التشويش Noise Models :

تم إجراء التجارب باستخدام كميات مختلفة من التشويش و بأنواع مختلفة من التشويش أيضاً. و بشكل عام، فإن صورة مشوهة يُمكن أن تُصاغ كما يلي:

$$X(i,j) = \begin{cases} O(i,j) & with\ a\ probability\ of\ 1-p \\ N(i,j) & with\ a\ probability\ of\ p \end{cases} \qquad (6\text{-}4)$$

حيث (i,j) هو موقع البكسل، و p النسبة المئوية لكمية التشويش و $N(i,j)$ قيمة التشويش و $O(i,j)$ هي قيمة البكسل الأصلية. و هناك بشكل رئيسي ثلاث أنواع من نماذج التشويش المُستخدمة في هذا البحث بالاعتماد على القيم التي يمكن أن تأخذها $N(i,j)$ و هي:

نموذج التشويش الثابت Fixed noise model: في هذا النموذج فإن $N(i,j)$ من أجل صورة L بت يمكن أن تأخذ قيمتين فقط $\{0,\ 2^L - 1\}$. من أجل صورة 8 بت فإن هاتين القيمتين للتشويش يمكن أن تأخذ 0 و 255.

نموذج التشويش العشوائي Random noise model: في هذا النموذج فإن $N(i,j)$ من أجل صورة L بت يمكن أن تأخذ أي قيمة في المجال $[0,\ 2^L - 1]$. من أجل صورة 8 بت فإن قيمة $N(i,j)$ تُختار بانتظام من المجال[0,255].

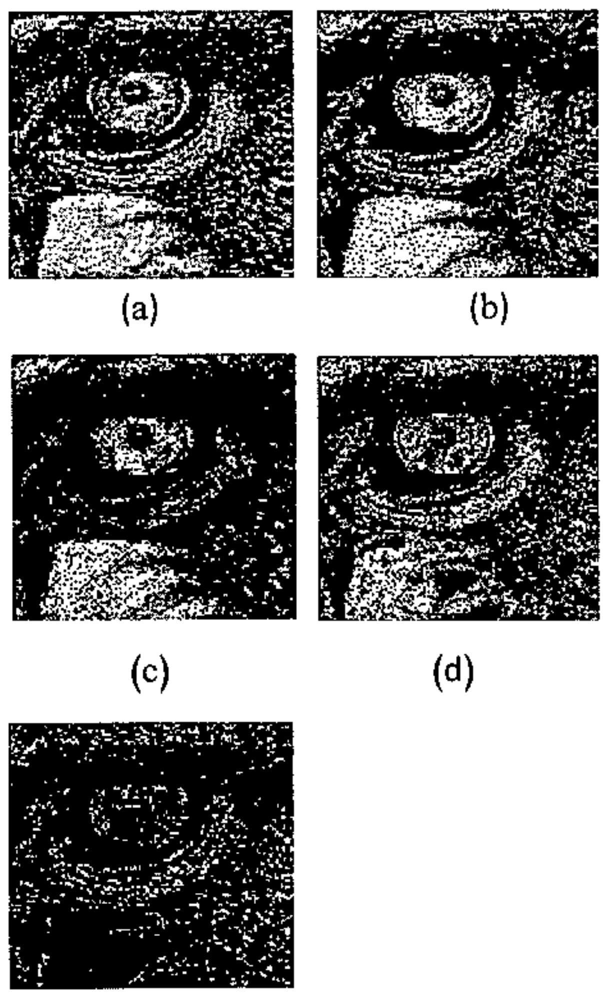
نموذج التشويش المضاف Additive noise model : في هذا النموذج فإن قيمة $N(i,j)$ تكون كما يلي:

$$N(i,j) = O(i,j) + X \qquad (7\text{-}4)$$

حيث X هي القيمة العشوائية المختارة بانتظام من المجال [0,255].

في النماذج الثلاث السابقة جميعها يتم اختيار بكسل عشوائي و يتم تشويه القناة العشوائية لذلك البكسل وفقاً لذلك.

يتم تطبيق نتيجة عينة من خوارزميات كشف التشويش على صورة قرد البابون المشوهة باستخدام 15 % من التشويش الثابت كما هو موضح في الشكل (4-3).



الشكل(4-3) نتائج خوارزميات كشف التشويش المختلفة (a) الجزء المُقرب من صورة البابون الأصلية (b) خوارزمية الكشف المُقترحة (c) مصفي المتجه التكيّفي[49] (d) ISAINF [60] (e) الطريقة في المرجع[40]

إن النتيجة المُبينة في الشكل (d)(4-3) مُولدة بقناع 7 × 7 بينما تستخدم البقية قناع 3 × 3 لكشف التشويش. إن البكسلات المشوهة تظهر في النتائج باللون الأحمر.

أُنجز التقييم النوعي للصورة المُستعادة و ذلك بإبراز الاختلاف بين الصورة الأصلية و المُستعادة.تم استخدام نسبة PSNR من أجل التقييم النوعي لنوعية الاستعادة.

$$\mathrm{PSNR} = 20\log_{10}\left(\frac{255}{\sqrt{MSE}}\right) \qquad (8\text{-}4)$$

$$\mathrm{MSE} = \frac{\sum_{i=1}^{N}\sum_{k=1}^{m}(x_{ik} - o_{ik})^2}{Nm} \qquad (9\text{-}4)$$

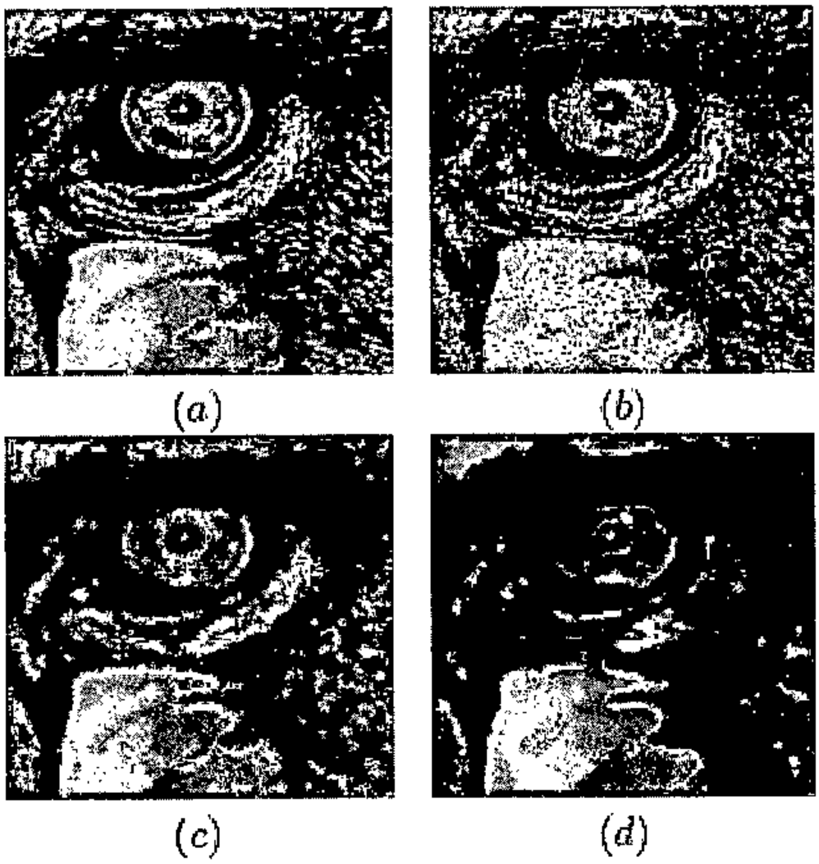
حيث m المجموع الكلي لمكونات اللون، و N هي المجموع الكلي لبكسلات الصورة و x_{ik} و o_{ik} المكون k من قناة تشويش بكسل الصورة و قيمته الأصلية في موقع البكسل "i" على الترتيب [61].

تم استخدام معيار " متوسط الخطأ المطلق" (Mean Absolute Error) أو MEA لتقييم قابلية المحافظة على التفاصيل و ذلك من أجل المُصفي المقترح.

$$\mathrm{MAE} = \frac{\sum_{i=1}^{N}\sum_{k=1}^{m}|x_{ik} - o_{ik}|}{Nm} \qquad (10\text{-}4)$$

يُستدل على الأداء المُتميز من خلال قيمة PSNR المرتفعة و قيمة MAE المنخفضة.

لتعليل الأداء المُتميز لخوارزمية تصفية التشويش المقترحة على VMF القياسية، نطبق كلا الطريقتين على صورة البابون المشوهة باستخدام 12% من التشويش العشوائي بدون كشف مُسبق للتشويش. و تم تقديم النتائج التجريبية في الشكل(4-4)[41].



الشكل(4-4) مقارنة لأداء تصفية التشويش بدون كشف التشويش على صورة البابون المشوهة (a) الصورة الأصلية (b) الصورة المشوهة (c) الطريقة المقترحة (d) VMF

يتجاوز أداء بعض التقنيات الضبابية أفضل الخوارزميات التي تعتمد على (-rank ordering). و نتيجة لذلك فقد قارنّا طريقتنا مع بعض التقنيات الضبابية.و تم تقديم النتائج في الشكل(4-6).

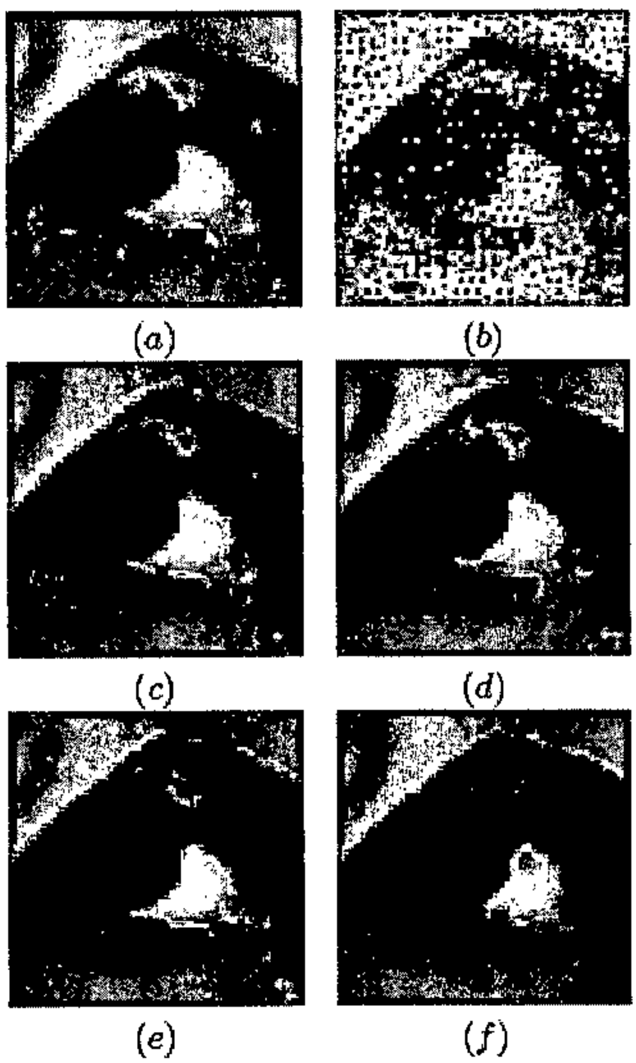
لقد قمنا بتجارب لمختلف مستويات ونماذج التشويش و على صور مختلفة، و ذلك لمقارنة أداء التقنية التي استخدمناها مع تقنيات مختلفة. و يُظهر الشكلان (4-5) و (4-7) نتائج المقارنة على صور الاختبار القِياسية Lena و الفلفل مّشوهة باستخدام 15% من التشويش التكيّفيّ و 5% من التشويش العشوائي على الترتيب.

تم توصيف الخوارزميات المختلفة من أجل الفلفل للحصول على الفرق بين الصور الأصلية و الصور المسترجعة. و أن قيمة كل بكسل في صورة الفرق هي الفرق المتجهي بين الصورة الأصلية و الصورة المسترجعة في فضاء اللون RGB. و نظراً لكون أغلب قيّم الفرق

تكمن في المجال[0,64]، فإن قيّم البكسلات في صور الفرق المُقدّمة تُمثل فقط القيم التي تكمن في المجال[0,64].

5-4 المناقشة Discussion :

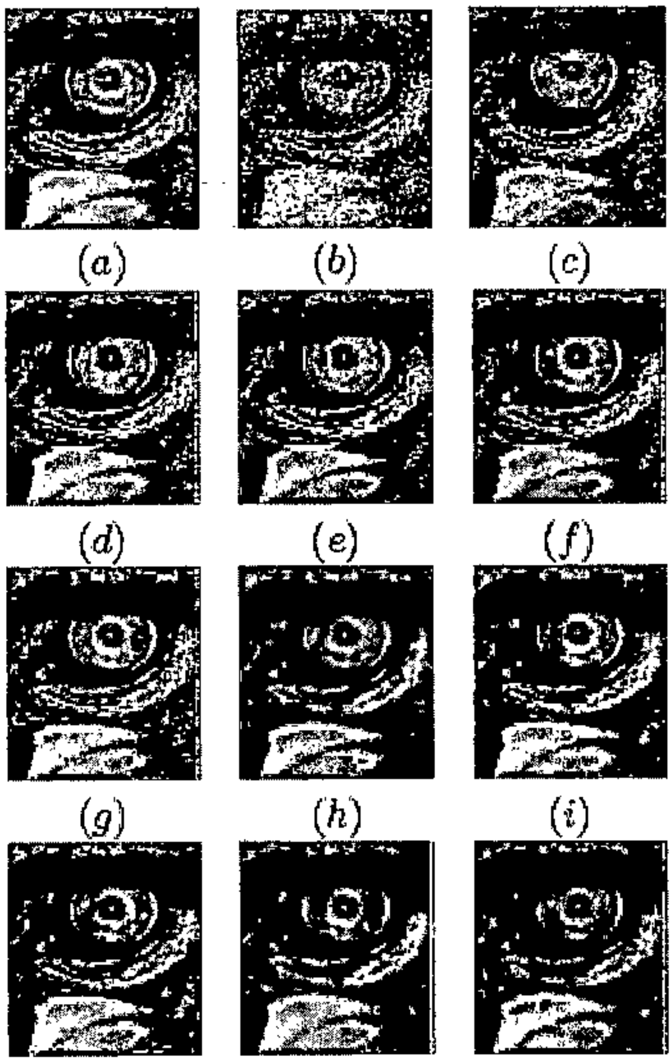
تُشير النتائج المُقدّمة في الشكل(4-3) إلى كفاءة HNN المُقترحة و المُعتمدة على خوارزمية الكشف مقارنة بخوارزميات الكشف القياسية و المستخدمة بشكل كبير. و كما تُشير النتائج إلى أن خوارزمية الكشف المُقترحة لها أقل نسبة خطأ إيجابي حتى على صورة البابون الكثيفة التفاصيل.



الشكل(4-5) مقارنة الخوارزميات المختلفة على صورة Lena القياسية (a) الجزر الأصلي المُقرب من الصورة (b) الصورة المُشوّهة (c)الخوارزمية المُقترحة (d) الطريقة في [40] (e) AMF [10] (f) VMF [41]

يُظهر الشكل(4-4) النتائج و التي تُبين بشكل واضح كفاءة الخوارزمية المُقترحة لتصفية التشويش على VMF القياسية و التي أعطت الصورة الكثير من الصقل.

في الشكل(4-5) تتم المقارنة بين خوارزميات مختلفة و ذلك بتطبيقها على صورة Lena القياسية من أجل 15% من التشويش التكيفيّ. و يُمكن الملاحظة بأن VMF تُقدم الكثير من الصقل بينما التقنية التي قُمنا باستخدامها تعمل بشكل جيد و تحافظ على التفاصيل و الحواف.



الشكل(4-6) (a) الصورة الأصلية، (b) الصورة المُشوّهة ،(c) الطريقة التي استخدمناها، FTSCF(d) ، FIDRM (e) ، DSFIRE (f) [57] ، PWLFIRE (g) [56] ، HAF (h) [71] ، LUM (i) [46] ، TSM (j) [66] ، MIFCF (k) [55] ، (l) TMAV [47]

و بملاحظة الشكل(4-6) نجد بأن التقنية التي قمنا باستخدامها يّمكن أن تُقارن بأفضل تقنية ضبابية ألا وهي FTSCF. و على الرغم من أنه يُمكن اعتبار طريقة FTSCF أفضل من التقنية التي استخدمناها، فإن لها عدة عيوب أولاً، تعتمد خوارزمية الكشف على دالة العضوية و التي تعتمد على ضبط عدد من الوسطاء. و يُمكن أن تتفاوت المجالات التي تنتمي لها قيّم هذه الوسطاء.كما أن طريقة إزالة التشويش ليست خطوة وحيدة و لكن تكرارية. بينما خوارزمية الكشف التي استخدمناها فإنها تعتمد على وسيط وحيد و الذي يأخذ قيمة وحيدة و يُعطي نتائج جيدة من أجل تلك القيمة و ذلك من أجل تشكيلة واسعة من التشويش و الصور المختلفة. و أن كلاً من الكشف و الإزالة ليسا بتكراريين و نتيجة لذلك فإن كلاً منهما يتميز بالسرعة و يُمكن أن يُستخدم في تطبيقات الزمن الحقيقي. تُظهر الصور المختلفة المُقدمة في الشكل(4-7) بوضوح تفوق الخوارزمية التي قمنا باستخدامها على الخوارزميات الأخرى.

الجدول(4-1) مقارنة قيّم MAE و PSNR من أجل 10% من التشويش الثابت المطبق على صورة Lena

الطريقة	MAE	PSNR
الصورة المشوهة	7.70	18.65
[20]FTSCF	0.37	51
طريقتنا المستخدمة	0.38	44.1
[11]AVFF	0.42	40.29
[15]MICM	.0.45	38.72
[26]SANRF	0.53	38.63
[24]FANRF	0.55	38.49
[12]AVLUMS	0.90	37.85
[23]FPGF	0.58	37.5
[10]AVMF	0.61	37.45
[16]VDOSF	0.72	34.72
[4]LIMF	2.32	34.38
[2]VMF	3.41	32.85

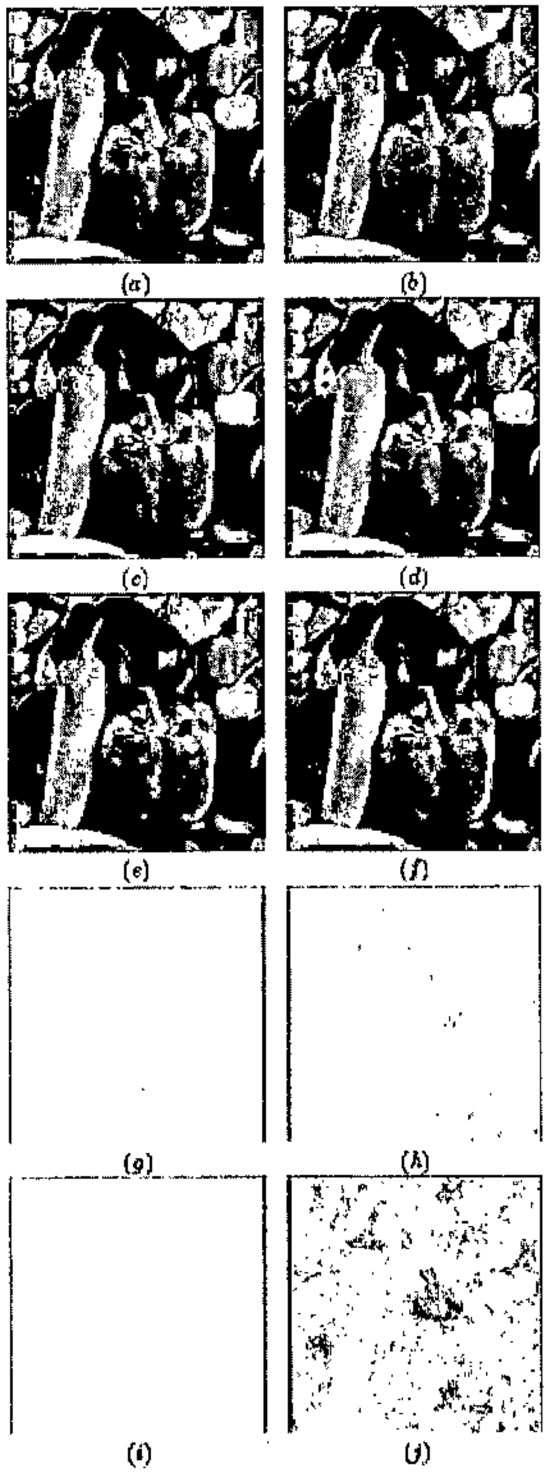
في الجدول(4-1) تتم مقارنة قيم PSNR و MAE لخوارزميات مختلفة من أجل 10% من التشويش الثابت على صورة Lena. تُظهر القيم الأداء المتفوق لطريقة FTSCF. وعلى الرغم من ذلك فإن أداء خوارزميتنا يُعتبر مثالياً حيث أنها تنقص MAE و تزيد PSNR. و لم يُشاهد هذا السلوك في خوارزميات مثل AVLUMS و التي تُظهر قيمة PSNR أكبر من AVMF و FPGF و VDOSF الخ و لكن لها قيمة أعلى من MAE.

6-4 الخلاصة Conclusion:

اقترحنا في هذا الفصل حل مكون من خطوتين للتشويش في الصور الملونة. و اعتمدت طريقة كشف التشويش على شبكة هوبفيلد العصبونية ثم طريقة تصفية التشويش و ذلك بالاستفادة من عيوب تقنية مصفي متوسط المتجه القياسية. و أن الحل المُقترح يُضاهي أفضل طرق التصفية المتوفرة و يُمكن أن يُطبق من أجل إزالة التشويش في الصور الطبيعية. و كما يُمكن استخدام الحل المقترح من أجل الصور ذات التدرج الرمادي أيضاً.

7-4 النظرة المستقبلية:

و كنظرة مستقبلية لهذا البحث يُمكن الاستفادة من تقنيات الذكاء الصنعي الأخرى مثل المنطق الضبابي و الخوارزميات الوراثية بالإضافة إلى الشبكات العصبونية و الدمج بينها للاستفادة من خواصها مجتمعة و ذلك لتصفية الصورة المشوشة لتصبح مماثلة تقريباً للصورة الأصلية و بأسرع وقت ممكن و بأقل كلفة حسابية


(a)
(b)
(c)
(d)
(e)
(f)
(g)
(h)
(i)
(j)

الشكل(4-7) نتائج المقارنة بين الخوارزميات المختلفة على صورة الفلفل المشوهة باستخدام 5% من التشويش العشوائي (a) الصورة الأصلية ،(b) الصورة المُشوّهة (c)، الطريقة التي استخدمناها ،(d) الطريقة في[40] ،(e) مصفي متوسط المتجه التكيّفي[49] ،(f)VMF[41] ،(g) صورة الفرق المكتسبة من خوارزميتنا،(h) صورة الفرق المكتسبة من[40] ،(i) صورة الفرق المكتسبة من [49] ،(j) صورة الفرق المكتسبة من VMF [41].

: References المراجع

[1] Zurada M. Jacek (1995) , "Introduction to Artificial Neural Systems" , PWS.

[2] عيسى، المهندس علام زكي، 2000-"الشبكات العصبية :البنية الهندسية، الخوارزميات"، التطبيقات. شعاع للنشر والعلوم ،الطبعة الأولى ،سوريا حلب.

[3] "الشبكات العصبونية الاصطناعية" من Wikipedia
http://ar.wikipedia.org/wiki/Artificial_neural_network

[4] "مقدمة في الشبكات العصبونية الاصطناعية " من Wikipedia
http://ar.wikipedia.org/wiki/Introduction_ To_Artificial_neural_network

[5] الشرقاوي، الأستاذ الدكتور محمد علي،" الذكاء الاصطناعي و الشبكات العصبية " مركز الذكاء الاصطناعي للحاسبات،الكتاب الأول.2001

[6] " الشبكة العصبونية Perceptron "
http://lessons.bdr130.net/2073.html

[7] K.C. Yow and R. Cipolla. "Feature-based human face detection". *Image and Vision Computing*, 15 (2004), pp. 713-735.

[8] T.F. Cootes and C.J. Taylor." Locating faces using statistical feature detectors". *Proceeding of the Second International Conference on Automatic Face and Gesture Recognition*, 2003, pp. 640-645.

[9] T.K. Leung, M.C. Burl, and P. Perona. " Finding faces in cluttered scenes using random labeled graph matching". *Proceedings of the Fifth International Conference on Computer Vision*, 2002, pp. 637-644.

[10] H.A. Rowley, S. Bluja, and T. Kanade. " Neural network-based face detection". *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 20 (1) (2005), pp. 23-38.

[11] K.K. Sung and T. Poggio. "Example-based learning for view-based human face detection". *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 20 (1) (2006), pp. 39-51.

[12] A.J. Colmenarez and T.S. Huang. "Face detection with information-based maximum discrimination". *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 2004, pp. 278-287.

[13] M. De Mariscoi, L. Cinque, and S. Levialdi. " Indexing pictorial documents by their content: a survey of current techniques". *Image and Vision Computing*, 15 (2004), pp. 119-141.

[14] B. Schiele and A. Waibel. " Gaze tracking based on face-color". Presented in International Workshop on Face and Gesture Recogntion, Zurich, July 2001.

[15] Y. Dai and Y. Nakano. " Face-texture model based on SGLD and its applications in face detection in a color scene". *Pattern Recognition*, 29 (6) (2002), pp. 1007-1017.

[16] J. Cai and A. Goshtasby. " Detecting human faces in color images". *Image and Vision Computing*, 18 (2006), pp. 63-75.

[17] Y. Miyake, H. Saitoh, H. Yaguchi, and N. Tsukada. " Facial pattern detection and color correction from television picture and newspaper printing". *Journal of Imaging Technology*, 16 (5) (2000), pp. 165-169.

[18] D. Androutsos, K.N. Plataniotois, and A.N. Venetsanopoulos. " A novel vector-based approach to color image retrieval using a vector angular-based distance measure". *Computer Vision and Image Understanding*, 75 (1/2) (July/August 2007), pp. 46-58.

[19] V. Rao and H. Rao. " C++ Neural Networks & Fuzzy Logic". MIS Press: New York, 2001.

[20] R. C. Gonzales, R. E. Woods, *Digital Image Processing*, Second Edition, Prentice-Hall, 2002.

[21] Veisi H., Jamzad M., Image Compression Using Neural Networks, Image Processing and Machine Vision Conference (MVIP), Tehran, Iran, 2005.

[22] N.Sonehara, M.Kawato, S.Miyake, K.Nakane, Image compression using
a neural network model, International Joint Conference on Neural Networks, Washington DC, 1989.

[23] G.L. Sicuranza, G. Ramponi, S. Marsi, Artificial neural network for image compression, Electronic letters 26, 477-479, 1990.

[24] S. Marsi, G. Ramponi, G. L. Sicuranza, Improved neural structure for
image compression, Proceeding of International Conference on Acoustic Speech and Signal Processing, Torento, 1991.

[25] S. Carrato, G. Ramponi, Improved structures based on neural networks
for image compression, IEEE Workshop on Neural Networks for Signal Processing, New Jersey, September 1991.

[26] S. Carrato, S. Marsi, Compression of subband-filterd images via neural
networks, IEEE Workshop on Neural Networks for Signal Processing. August 1992.

[27] G. Qiu, M. Varley, T. Terrel, Image compression by edge pattern learning using multilayer perceptron, Electronic letters, Vol 29, No 7, April 1993.

[28] R.Sentiono, G. Lu, Image compression using a feedforward neural network, International Conference on Neural Networks, 1994.
[29] J. Jiang, Image compression with neural networks -A survey, Image Communication, ELSEVIER, Vol. 14, No. 9, 1999.

[29] C. Cramer, Neural networks for image and video compression: A review, European Journal of Operational Research, Vol. 108, July 1998.

[30] B. Verma, M. Blumenstein, and S. Kulkarni, A Neural Network Based
Technique for Data Compression, Proceedings of the IASTED International Conference on Modelling and Simulation, MSO97, Singapore, 1997.

[31] A.Namphol, S.Chin, M. Arozullah, Image compression with a hierarchical neural network, IEEE Trans. Aerospace Electronic Systems Vol. 32 No.1, January 1996.

[32] J. S. Lin, S.H. Liu, A competitive continuous Hopfield neural network
for vector quantization in image compression, Engineering Applications of Artificial Intelligence, Vol. 12, 1999.

[33] G. Pavlidis, A. Tsompanopoulos, A. Atsalakis, N. Papamarkos, C. Chamzas, A Vector Quantization – Entropy Coder Image Compression System, IX Spanish Symposium on Pattern Recognition and Image Processing, 2001.

[34] C. Amerijckx, J. D. Legaty, M. Verleysenz, Image Compression Using
Self-Organizing Maps, Systems Analysis Modeling Simulation Vol. 43, No. 11, November 2003.

[35] S. Costa, S. Fiori, Image compression using principal component neural
networks, Image and vision computing, Vol. 19, 2001.

[36] M. Egmont-Petersen, D. de Ridder, and H. Handels, Image processing
with neural networks - a review, Pattern Recognition, vol. 35, pp. 2279-2301, 2002.

[37] A. Rahman and Chowdhury Mofizur Rahman, "A New Approach for Compressing Color Images using Neural Network", Proceedings of International Conference on Computational Intelligence for Modeling, Control and Automation - CIMCA 2003 , Vienna, Austria, 2003.

[38] Thomas M. Cover and Joy A. Thomas. Elements of Information Theory,
John Wiley and Sons Inc., New York, N.Y., 1991.

[39] Sonja Grgic, Marta Mrak, Mislav Grgic, Comparison of JPEG Image Coders Proceedings of the 3rd International Symposium on Video Processing and Multimedia Communications, VIPromCom, pp. 79-85, Zadar, June 2001.

[40] N. Alajlana, M. Kamela, and E. Jernigan. Detail preserving impulsive noise removal. *Signal Processing: Image Communication*, 19(10):993–1003, November 2004.

[41] J. Astola, P. Haavisto, and Y. Neuvo. Vector median filters. *Proceedings of the IEEE*, 78(4):678–689, Apr 1990.

[42] M. Bartkowiak and M. Domanski. Vector median filters for processing of color images in various color spaces. *Fifth International Conference on Image Processing and its Applications*, pages 833–836, Jul 1995.

[43] A. Beghdadi and A. Khellaf. A noise-filtering method using a local information measure. *IEEE Transactions on Image Processing*, 6(6):879–882, Jun 1997.

[44] S. Chartier and R. Lepage. Learning and extracting edges from images by a modified hopfield neural network. *16th International Conference on Pattern Recognition*, 3:431–434
vol.3, 2002.

[45] F.Farbiz and M.B.Menhaj. A fuzzy logic control based approach for image filtering. In *Fuzzy Techniques in Image Processing*, volume 52, pages 194–221, Heidelberg, Germany,

2000. Physica Verlag.

[46] R. Hardie and C. Boncelet. Lum filters: A class of rank-order based filters for smoothing and sharpening. *IEEE Transactions on Image Processing*, 41(3):1061–1076, March 1993.

[47] H. K. Kwan and Y. Cai. Fuzzy filters for image filtering. *45th Midwest Symposium on Circuits and Systems*, 3(12):672–675, 2002.

[48] X. Liu, B. Li, Z. Su, X. Shi, and F. Liu. Impulsive noise removal by nonlocal fuzzy diffusion. *Fourth International Conference on Image and Graphics*, pages 155–159, Aug. 2007.

[49] R. Lukac. Adaptive vector median filtering. *Pattern Recognition Letters*, 24(12):1889–1899, 2003.

[50] R. Lukac, V. Fischer, G. Motyl, and M. Drutarovsky. Adaptive video filtering framework. *IJIST*, 14(6):223–237, 2004.

[51] R. Lukac and S. Marchevsky. Adaptive vector lum smoother. *International Conference on Image Processing*, 1:878–881 vol.1, 2001.

[52] W. Luo and D. Dang. An efficient method for the removal of impulse noise. *IEEE International Conference on Image Processing*, pages 2601–2604, Oct. 2006.

[53] K. Manglem Singh and P. Bora. Adaptive vector median filter for removal impulses from color images. *ISCAS '03. Proceedings of the International Symposium on Circuits and Systems*, 2:II–396–II–399 vol.2, May 2003.

[54] J. Park and L. Kurz. Image enhancement using the modified icm method. *IEEE Transactions on Image Processing*, 5(5):765–771, May 1996.

[55] R.Lukac. Color image filtering by vector directional order statistics. *Pattern Recognition and Image Analysis*, pages 279–285, 2002.

[56] F. Russo. Fire operators for image processing. *Fuzzy Sets and Systems*, 103(2):265–275, April 1999.

[57] F. Russo and G. Rumponi. Removal of impulse noise using a fire filter. *IEEE Transactions on Image Processing*, 1:975–978, 1996.

[58] S. Schulte, V. De Witte, M. Nachtegael, D. Van Der Weken, and E. Kerre. A fuzzy impulse noise detection and reduction method. *IEEE Transactions on Image Processing*, 15(5), 2006.

[59] S. Schulte, V. De Witte, M. Nachtegael, D. Van Der Weken, and E. Kerre. Fuzzy two-step filter for impulse noise reduction from color images. *IEEE Transactions on Image Processing*, 15(11):3567–3578, Nov. 2006.

[60] F.-l. C. Shitong Wang, Yueyang Li and M. Xu. An iterative self-adaptive algorithm to impulse noise filtering for color images. *International Journal of Information Technology*, 11(10), 2005.

[61] B. Smolka. On the adaptive impulsive noise attenuation in color images. *Image Analysis and Recognition*, 19(10):307–317, 2006.

[62] B. Smolka and A. Chydzinski. Fast detection and impulsive noise removal in color images. *Real-Time Imaging*, 11(5-6):389–402, 2005.

[63] B. Smolka, R. Lukac, A. Chydzinski, K. N. Plataniotis, and W. Wojciechowski. Fast adaptive similarity based impulsive noise reduction filter. *Real-Time Imaging*, 9(4):261–276, 2003.

[64] B. Smolka, R. Lukac, and K. Plataniotis. Fast noise reduction in cdna microarray images. *23rd Biennial Symposium on Communications*, pages 348–351, 29 - June 1, 2006.

[65] B. Smolka, K. Plataniotis, A. Chydzinski, M. Szczepanski, A. Venetsanopoulos, and K. Wojciechowski. Self-adaptive algorithm of impulsive noise reduction in color images. *Pattern Recognition*, 35(8):1771–1784, August 2002.

[66] K. M. T.Chen and L. Chen. Tri-state median filter for image denoising. *IEEE Transactions on Image Processing*, 8(12):1834–1838, December 1999.

[67] P. Trahanias, D. Karakos, and A. Venetsanopoulos. Directional processing of color images: theory and experimental results. *IEEE Transactions on Image Processing*, 5(6):868–880, Jun 1996.

[68] K. Valavanis, J. Zheng, and J. Gauch. On impulse noise removal in color images. *IEEE International Conference on Robotics and Automation*, pages 144–149 vol.1, Apr 1991.

[69] T. Viero, K. Oistamo, and Y. Neuvo. Three-dimensional median-related filters for color image sequence filtering. *IEEE Transactions on Circuits and Systems for Video Technology*, 4(2):129–142, 208–10, Apr 1994.

[70] Y. Wakabayashi and A. Taguchi. A new efficient approach for removal of impulse noise for color images. *APCCAS 2006. IEEE Asia Pacific Conference on Circuits and Systems*, pages 85–88, Dec. 2006.

[71] J.-H.Wang and H.-C. Chiu. Haf: An adaptive fuzzy filter for restoring highly corrupted images by histogram estimation. *Proc. Natl. Sci. Counc. ROC(A)*, 23(5):630–643, April 1999.

[72] M. Yu, G. Jiang, and B. Yu. Noise detection based impulse noise removal for color image. *The 2000 IEEE Asia-Pacific Conference on Circuits and Systems*, pages 453–456, 2000.

[73] P. Zvonarev, I. Apalkov, V. Khryashchev, and I. Reznikova. Neural network adaptive switching median filter for the restoration of impulse noise corrupted images. In *ICIAR05*, pages 223–230, 2005.

overlapping the blocks is allowed and choosing the networks in compressor is based on the Best-SNR.

In additional, this search includes a method for noise removal in corrupted color images is discussed. The new method consists of two phases. The first phase is a noise detection phase where a modified Hopfield neural network is used to detect noise pixels. The second is a noise filtering phase where the disadvantage of taking Vector Median in a single color space is addressed and a new algorithm based on performing Vector Median first in RGB space and then in HSI space is presented. The results of simulations performed on a set of standard test images on a wide range of noise corruption show that the proposed method is capable of detecting all the impulse noise pixels with almost zero false positive rates and removes noise while retaining finer image details.

The chapters of search also includes a theoretic introduction about the neural networks in this search, learning method, and training algorithms and the future horizons.

Abstract

٦٩١٣٢٤

Recently, the technical difficulties has increased and become more complexity and the traditional solutions programmatic become incapable of realizing their, that led the researchers to understanding deeply the mechanism of the human thinking and how it processes, store , and recall information and depending on the simulation manner to solve those problems that was performed to associated skeletal for programmatic structure called the Artificial Neural Networks.

This research clears the capability of neural networks in problems simulation in different domains like digital images processing domain by results comparing with traditional processing methods.

This research includes three neural network-based systems building for detecting human faces in color images. The three methods vary in the type of input fed into the network. We discuss networks based on (1) interpolated images, (2) RGB histograms, and (3) YES histograms. Each system was trained on the same image set using the Levenberg-Marquad method. Computational results for the training set and a common test image set are presented.

The research also includes an adaptive method for image compression that is based on complexity level of the image. The basic compressor/de-compressor structure of this method is a multilayer perceptron artificial neural network. In adaptive approach different Back-Propagation artificial neural networks are used as compressor and de-compressor and this is done by dividing the image into blocks, computing the complexity of each block and then selecting one network for each block according to its complexity value. Three complexity measure methods, called Entropy, Activity and Pattern-based are used to determine the level of complexity in image blocks and their ability in complexity estimation are evaluated and compared. In training and evaluation, each image block is assigned to a network based on its complexity value. Best-SNR is another alternative in selecting compressor network for image blocks in evolution phase which chooses one of the trained networks such that results best SNR in compressing the input image block. In our evaluations, best results are obtained when